A Multi-Criteria Methodology for developing a Sustainability Assessment Tool

Moeez Ahmed

Thesis to obtain the Master of Science Degree in

Energy Engineering and Management

Supervisors

Teresa Sofia Cipriano Gonçalves Rodrigues

Pedro Miguel Mateus Carreira

Examination Committee

Chairperson: Susana Isabel Carvalho Relvas

Supervisor: Pedro Miguel Mateus Carreira

Member: Ana Catarina Lopes Vieira Godinho de Matos

November 2019

Acknowledgements

To Professor Teresa Rodrigues, for all the support, patience and dedication.

To Pedro Carreira, and to PLANETIERS, for their continued encouragement and assistance.

To my friends, for the continued motivation and good mood.

To my family, for always being there for me.

Thank you.

Abstract

Excessive exploitation of natural resources has resulted in the depletion of natural reserves. The increase in the awareness of natural resource depletion has led to a drastic increase in the demand for sustainable items. No particular system is present to assess sustainability of products leading to incoherency and ad hoc evaluations which do not incorporate all potential criteria pertinent for examination. Although extensive literature is available, all of them fail to account assessing sustainability of consumer products.

The work performed in thesis aimed to develop a methodological framework with a purpose of creating a tool to assess sustainability of products while solving a real-world problem and aiming to contribute to the literature. The proposed framework combines the principles of Multi Criteria Decision Analysis (MCDA) and Participatory Processes. The identification of criteria was achieved by employing, Welphi, a web-Delphi platform. The tool was constructed through a socio-technical interaction between the facilitator and the decision makers by implementing MACBETH approach for the creation an evaluation model aided by M-MACBETH, a decision support system.

The framework was implemented with PLANETIERS as the case study to build the assessment tool that enabled the decision makers to assess performance of the products and evaluate their overall sustainability. The results obtained provided valuable addition to the literature regards to socio-technical interactions in sustainability assessment and employment of M-MACBETH and Welphi in decision aiding and participatory process. The methodology implemented has proved to be robust and allows forth adaptation towards products from various categories.

Keywords: Sustainability Assessment, MCDA, Participatory Processes, MACBETH, Delphi

Resumo

A exploração excessiva de recursos naturais resultou no esgotamento de reservas naturais. O crescimento da consciencialização da escassez dos recursos naturais causou um aumento drástico na procura de bens de consumo sustentáveis. Não existe nenhum sistema de avaliação da sustentabilidade de produtos, causando incoerências e avaliações que não incorporam todos os potenciais critérios pertinentes á analise. Apesar da existência de literatura extensa no assunto, nenhum dos trabalhos anteriores toma em consideração a avaliação da sustentabilidade de produtos de consumo.

A presente tese visa desenvolver de uma estrutura metódica com o propósito de criar uma ferramenta para avaliar a sustentabilidade de produtos, resolvendo um problema real e contribuindo para a literatura existente. A estrutura proposta combina os princípios de analise de decisão multicritério (MCDA) e processos participativos. A identificação dos critérios foi feita através da plataforma Welphi, uma plataforma Delphi online. A ferramenta foi construída numa relação sociotécnica entre o promotor e os especialistas, implementando o método MACBETH para a criação de um modelo de avaliação no programa M-MACBETH, um sistema de suporte de decisão.

A estrutura foi implementada com a empresa PLANETIERS num caso de estudo para construir uma ferramenta de avaliação que permitiu que especialistas avaliassem a sustentabilidade de produtos. Os resultados obtidos são uma adição importante á literatura existente graças às interações sociotécnicas usadas na avaliação da sustentabilidade e ao uso do M-MACBETH e Welphi no processo de decisão e processo participativo. A metodologia demonstrou ser robusta e poderá ser adaptada a produtos de várias categorias.

Palavras-Chave: Avaliação de Sustentabilidade, MCDA, Processos Participativos, MACBETH, Delphi

Table of Contents

Acknowledgements ii

Abstract iii

Resumo iv

List of Figures vii

List of Tables ix

1 Introduction 1

1.1 Thesis Objective 2

1.2 Thesis Outline 3

2 Literature Review 4

2.1 Sustainability 4

2.1.1 Sustainability Assessment 6

2.2 Assessment Methods 7

2.2.1 Multi Criteria Decision Analysis 8

2.3 Participatory Methods 13

2.3.1 Delphi 15

3 Framework Methodology 18

3.1 Outline 18

3.2 Problem Context 20

3.3 Identifying Criteria 21

3.4 Web Delphi 21

3.4.1 Panel 22

3.5 Structuring Phase 22

3.5.1 Value Tree 22

3.5.2 Descriptors 23

3.5.3 Evaluation Phase 24

3.5.4 Criteria Weighting Procedure 26

3.5.5 Additive Aggregate Model 27

4 Framework Application 29

4.1 Generation of Potential Criteria 29

4.2 Delphi Process 30

4.2.1 Delphi Panel 32

4.2.2 Delphi Application 32

4.3 Model Implementation 35

4.3.1 Screening Criterion 36

4.3.2 Value Trees 36
4.3.3 Descriptors of Performance 38
4.3.4 Value Functions 46
4.3.5 Weighting Criteria 47
4.3.6 Category Thresholds 49
5 Results 53
5.1 Web Delphi Results 53
5.1.1 Participants Response Rate 53
5.1.2 Criterion Agreement 54
5.2 Evaluation Model Results 59
5.2.1 Reusable Model Results 59
6 Discussion 65
6.1 Delphi Survey 65
6.1.1 Welphi Contribution 65
6.1.2 Delphi Response 65
6.1.3 Panel 65
6.2 Multicriteria Model 66
6.2.1 MACBETH Approach 66
6.2.2 Structuring Value Tree 66
6.2.3 Generation of Value Function 66
6.2.4 Determination of Criteria Weights 67
6.2.5 Supplier Form and Product Classification 67
6.2.6 Generic and Category Specific Model 67
6.2.7 Decision Support System 68
6.2.8 Limitations 68
6.2.9 Contribution to Literature 69
6.2.10 Possible Applications and Future Work 69
7 Conclusions and Future Work 70
References 72
Appendix A 78
Appendix B 81
Appendix C 85
Appendix D 89

List of Figures

Figure 1 - Framework for Development of the Assessment Tool 19
Figure 2 - Operations Schematic 20
Figure 3 - Judgement Matrix (Adapted from Bana e Costa et al, 2008, "Development of Reusable Bid Evaluation Models for the Portuguese Electric Transmission Company") 25
Figure 4 - Judgement Matrix (Adapted from Bana e Costa et al, 2012, "MACBETH") 26
Figure 5 - Web-Delphi Process Flow Diagram 31
Figure 6 - Screenshot of 3rd Round 33
Figure 7 - Agreement Protocol 35
Figure 8 - Generic Value Tree - "Product Sustainability Assessment Tool" 37
Figure 9 - Descriptors of Performance for the criterion "Product Disposability" 39
Figure 10 - Judgement Matrix for the "Product Disposability" 46
Figure 11 - Graphical representation of the value function for the criterion "Ratio of Waste Produced" 47
Figure 12 - Generic Model Weights "Product Sustainability Assessment Tool" 49
Figure 13 - Bottom-Up procedure to identify reference profile 50
Figure 14 - Category threshold Scheme 51
Figure 15 – Web Delphi Response Rate 53
Figure 16 - Indicator Approval with respect to Areas of Concern 55
Figure 17 - Reusable Model Results 59
Figure 18 - Weighted Differences Profiles, "Plate vs Bag 1" 60
Figure 19- Weighted Differences profiles, "Straw vs Bottle 1" 61
Figure 20 - XY Map, "Product Repair vs Welfare Impact" 61
Figure 21 - Sensitivity Analysis "Welfare Impact" 62
Figure 22 - Robustness Analysis in absence of Cardinal Global Information 63
Figure 23 - Robustness Analysis with Cardinal Global Information 63
Figure 24 - Robustness Analysis with Uncertainty 64
Figure 25 - Delphi 1st Round Invitation Email 78
Figure 26 - Delphi Invitation Email 2nd round 79
Figure 27 - Delphi Reminder Email 80
Figure 28- Delphi Last Reminder 80
Figure 29 - Value Tree Reusables 81
Figure 30 - Value Tree Fashion 82
Figure 31 - Value Tree Home Décor 83
Figure 32 - Value Tree Hygiene 84
Figure 33 - Value Scale Recycled Materials in Product 85
Figure 34 - Value Scale Renewable Materials in Product 85
Figure 35 – Value Scale Organic Farming 85
Figure 36 - Value Scale Recyclability of Waste 85
Figure 37 - Value Scale Renewable Energy 85
Figure 38 - Value Scale Biodegradable 85
Figure 39 - Value Scale Product Recyclability 86
Figure 40 - Value Scale Product Repairability 86
Figure 41 – Value Scale Product Disassembly 86
Figure 42 - Value Scale Product Disposability 86
Figure 43 - Value Scale Product Packaging 86
Figure 44 - Value Scale Eco-Label 86
Figure 45 - Value Scale CO2 Sequestration 87

Figure 46 - Value Scale Green Delivery 87
Figure 47 – Value Scale Ecological Mitigation Effort 87
Figure 48 - Value Scale ISO 14000 Certification 87
Figure 49 - Value Scale Welfare Impact 87
Figure 50 - Value Scale Locally Employed 87
Figure 51 – Value Scale Animal Welfare 88
Figure 52 - Value Scale Water Consumed 88
Figure 53 – Value Scale Energy Consumed 88
Figure 54 - Value Scale Electricity Consumed by Product 88
Figure 55 - Value Scale Ratio of Waste Produced 88
Figure 56 - Overall Thermometer Fashion 89
Figure 57 - Overall Thermometer Hygiene 89
Figure 58 - Overall Thermometer Home Decor 89

List of Tables

Table 1 - PLANETIERS' Principles 29
Table 2 - Delphi Schedule 34
Table 3 - List of Criteria and their descriptions 39
Table 4 - Constructed Performance Scale for "Recycled Materials in Product" (Environmental) 40
Table 5 - Constructed Performance Scale for "Renewable Materials in Product" (Environmental) 40
Table 6 - Constructed Performance Scale for "Organic Farming" (Environmental) 41
Table 7 - Constructed Performance Scale for "Recyclability of Waste" (Environmental) 41
Table 8 - Constructed Performance Scale for "Renewable Energy" (Environmental) 41
Table 9 - Constructed Performance Scale for "Biodegradable" (Environmental) 41
Table 10 - Constructed Performance Scale for "Product Recyclability " (Environmental) 41
Table 11 - Constructed Performance Scale for "Product Repairability " (Environmental) 42
Table 12 - Constructed Performance Scale for "Eco-Label " (Environmental) 42
Table 13 - Constructed Performance Scale for "Product Disassembly " (Environmental) 42
Table 14 - Constructed Performance Scale for "Product Disposability " (Environmental) 42
Table 15 - Constructed Performance Scale for "Product Packaging" (Environmental) 43
Table 16 - Constructed Performance Scale for "Green Delivery" (Environmental) 43
Table 17 - Constructed Performance Scale for "CO2 Sequestration" (Environmental) 43
Table 18 - Constructed Performance Scale for "Ecological Mitigation Effort" (Environmental) 44
Table 19 - Constructed Performance Scale for "ISO 14000 Certification of Intermediary Suppliers " (Environmental) 44
Table 20 - Constructed Performance Scale for "Welfare Impact" (Social) 44
Table 21 - Constructed Performance Scale for "Locally Employed" (Social) 44
Table 22 - Constructed Performance Scale for "Animal Welfare" (Social) 45
Table 23 - Quantitative Performance Scale for "Ratio of Waste Produced" (Economical) 45
Table 24 - Constructed Performance Scale for "Water Consumed" (Economical) 45
Table 25 - Constructed Performance Scale for "Energy Consumed" (Economical) 45
Table 26 - Constructed Performance Scale for "Electricity Consumed by Product" (Economical) 45
Table 27 - Product Category Specific Model Weights 51
Table 27 - Product Category Specific Model Weights (Continued) 52
Table 28 - Indicator Proposed, Approved and Rejected by Areas of Concern and Dimensions 54
Table 30 - Aggregated Responses on each indicator by area of concerns and dimension 56

1 Introduction

An ecosystem is a complex interconnected self-sustaining system of living and inanimate organisms and the excessive exploitation of these natural resources has caused an irreparable damage to the ecosystem. Over consumption of resources has depleted the capacity of our environment to replenish its natural reserves resulting in an ecological overshoot where humanity uses an equivalent of 1.7 Earths to maintain its current living patterns (GFN, 2019). Coupled with the increasing population and growing economies, henceforth, depleting the reserves for a normative existence for future generations. It has also inhibited the capacity of the natural system to replace the anthropogenic induced carbon dioxide thus accelerating the process of global warming in line with linear consumption of raw materials combined with fossil fuel power generation, has led to the increase in average global temperatures triggering a change in Earth's climate (NRDC, 2019).

Citizens are influencing national policy to adopt measures to combat climate change and instate laws to protect the environment. Corporations are active in climate change and management activities when they are closer to consumers (Haddock-Fraser and Tourelle, 2010). Reputation in the eyes of the customer is an important factor for corporations even if they do not achieve cost-reduction benefits (Chapman and Shigetomi, 2018). Public demand is pressurizing companies and governments to administer sustainable practices and policies and embody a more ecofriendly and "green" stance. Market stakeholder influences is one factor behind corporate environmental responsibility practices and is more pronounced in developed countries (Dogl and Behnam, 2014). Individual and collective decision making from daily consumptions to long term investments can impact the natural environment. Development of new policies can incentivize individuals, households and firms to make environmentally sustainable decisions (OECD, 2017). European Union has set targets to cut greenhouse gas emissions by 40% compared with 1990 and increase in energy efficiency of at least 32.5% as per the 2030 Key European Union (EU) targets (European Commission, 2014). These targets do not necessarily involve changing policies towards renewable energies or energy efficiency. Behavioral changes and societal awareness within the community can bring about the realization of these targets.

Global warming, climate change and unsustainable consumption of resources has now been acknowledged as the leading concerning matter in the 21st century and large inroads have been made towards a sustainable lifestyle on domestic, commercial and industrial levels in Europe and abroad. New legislation has been implemented to regulate industries, multinational firms, commercial and domestic infrastructure. The EU has significantly contributed in promoting sustainable activities and conservation of natural resource. However, in regards to everyday commodities and household consumables a standardized tool for assessing their sustainability has not been developed. Being cognizant of the demand of ecofriendly products, PLANETIERS has felt the need to plug the gap and standardize its commercial process.

PLANETIERS is an e-commerce platform exclusively marketing environmentally friendly products based in Lisbon (PLANETIERS, 2019). PLANETIERS aims to bridge the gap between environmentally conscious individuals and sustainability driven companies. As more citizens are inclining towards buying environmentally friendly products, PLANETIERS intends to brand its platform as the leading market for low carbon environmental goods and services. Their aim is to provide a holistic solution towards sustainable development in a single platform. Their operational structure is based on three business units:

- Online Sustainable Marketplace
- Blogs and News
- Events

PLANETIERS is continuously engaging with suppliers and receive numerous products on a regular basis. It is paramount for PLANETIERS to accept products that are sustainably produced and for that reason it is mandatory for them to distinguish if the products satisfy their standards to be allocated into their platform. PLANETIERS encompasses an extremely diverse product portfolio in its database, which demands a manual inspection of the product and validity of its sustainability. In order to ascertain their sustainability, the products are evaluated on an individual basis. The resultant product analysis determines if the product is sustainable (or not) and is then allocated to their online platform (or rejected if it does not meet their standards). The manual evaluation of products emanating from a range of categories often requiring distinct parameters for assessing sustainability, slows down the operation considerably and is more opportune to human error. As their online platform is growing and incorporating more merchandize into their database, manual inspection of items is time-consuming and impractical. PLANETIERS is upgrading their platform and wish to automize the appraisal process by allowing suppliers to fill in the product specifications online and this information in turn will be utilized to assess the products sustainability. PLANETIERS intends to implement a system with an aim to ease the evaluation process and support the appraisal of their inventory with greater efficiency, accuracy and reliability, and incorporating greater deal of flexibility by standardizing the process and allowing the assessment of sustainability for distinct products under a single platform. In order to achieve this objective, a robust and reliable tool needs to be developed that can evaluate their diverse range of products.

1.1 Thesis Objective

The scope of the thesis pertains to identifying sustainable products within their inventory as well as appraising additional products arising from induction of new suppliers, by analyzing their database, and categorizing and ranking their product portfolio with respect to sustainability. For this purpose, the development of an assessment tool was proposed. This dissertation aims to present the work performed by the author, in collaboration with PLANETIERS and Instituto Superior Tecnico (IST), concerning the development of a tool to be used by the PLANETIERS. In order to develop the assessment tool, a multicriteria analysis model needs to be established. To achieve this realization, a methodological framework is designed which will be elaborated upon later in this thesis. The focus of study in this dissertation is sequenced in two parts; 1) the employment of Delphi methodology through a web-based tool WELPHI, a platform for participatory processes in decision-making, and 2) conducting a decision conference in pursuance of establishing the product appraisal tool, grounded on the MACBETH methodology, facilitated by the M-MACBETH decision support system. The aforementioned set of activities within the multicriteria framework will capacitate the PLANETIERS' organization to discern if a product is eligible to be permitted into their online platform. Moreover, the process of construction and implementation of the proposed framework is by itself an opportunity for improvement, since it is intended to generate awareness and discussion within the relevant stakeholders and ultimately create a transparent, user-friendly and well-understood decision support tool.

1.2 Thesis Outline

The first and present chapter aims to provide the underlying objectives and purposes for the development of this thesis. The literature review is presented in the subsequent chapter, detailing the concept of sustainability and examining the current assessment methods for evaluating sustainability and participatory methods for socio-technical process. The remainder of this dissertation paper is organized as follows: Section 2 represents the literature review while framing the decision problem. Section 3 elaborates on the design of the framework and Section 4 depicts the implementation of the Delphi and MACBETH methodologies. Results from the application of the Delphi process and the multicriteria model will be presented in Section 5. Discussion with regards to this thesis study is conveyed in Section 6 and finally, Conclusions will be provided in Section 7.

2 Literature Review

As previously described, PLANETIERS is currently characterized by the situation of not having an established system of gauging the sustainability of the products, present on their online repository and of receiving additional products, culminating in non-conformity in the assessment of their merchandise. This facilitates the need for creation of a product appraisal tool which in turn requires the development of methodology with social and technical components. The technical component involves the model building process and the social aspect elicits the expertise of the specialists and are expanded upon in the assessment methods and participatory process respectively. This chapter is divided into three parts:

- Concept of sustainability
- Assessment methods
- Participatory processes

Despite the bevy of existent methodologies that address sustainability assessment in various contexts, the consulted literature did not provide a comprehensive resolution for the PLANETIERS case study and the task has expanded towards developing a framework for this purpose. Nevertheless, the existing literature provides the conceptual foundations necessary for the establishment of a new framework encompassing features of the particular subject matter.

The literature review was mainly based on the literature available on the “ScienceDirect” and “Google Scholar” online databases. These databases were searched using the queries: “sustainability evaluation methodologies”, “sustainability assessment of household commodities”, “evaluating sustainability in small and medium enterprises”, “multicriteria decision analysis in sustainability” and “indicators for measuring sustainability”. The resultant outcomes of the keywords were searched and relevant topics were filtered based on the article title. Once a relevant paper was chosen, its abstract and conclusions was studied. If the selected paper befitted the context of this dissertation it was further deliberated.

2.1 Sustainability

The world population is growing at an unprecedented rate with the expected rise in population to 8.6 billion in 2030 and 9.8 billion by 2050 according to the World Population Prospects: The 2017 Revision, published by the UN Department of Economic and Social Affairs (DESA, 2017). This entails an increasing stress on the world’s resources which is currently unable to regenerate its resources at a similar rate it is being consumed. This inefficient use of resources has led to unprecedented levels of waste generation. An estimate of 2.01 billion tons of municipal solid waste is generated annually with at least 33 percent is not managed in an environmentally safe manner and global waste is expected to grow to 3.40 billion tons by 2050, more than double the population growth over the same time period (WB, 2019). Waste disposal has become an increasingly prominent issue in the context of environmental protection. Environmental degradation, scarcity of resources and societal issues have reshaped agendas and governance, with a strong emphasis on adoption of “Green” strategies.

The definition of sustainability which caters the perspective of the PLANETIER’s organization was proposed in the 1994’s Oslo Symposium on Sustainable Consumption (NMCE, 1994) stating the use of products and services which meet our basic needs or bring a better quality of life while minimizing the use of natural resources and toxic materials with reduced emissions of waste and pollutants over the life cycle of the product or service in order to protect the needs of the future generations. This definition allies with PLANETIERS’s objectives of fostering a culture of an ecofriendly lifestyle and propagating the use of eco-innovative products and services. To pursue sustainability is to create and maintain the conditions under which humans and nature can exist in productive harmony to support present and future generations (EPA, 2016). The products will be environmentally friendly throughout their lifetime and will pose no permanent damage to the environment from resource extraction, product manufacture and its final disposal. Its production intends to increase efficiency and reduce resource consumption alongside minimizing waste generation and curb pollutant emissions, while employing use of renewable resources and energy.

Sustainability is an ambiguous term which is used in a broader context or for a specific domain (cultural, urbanization, economy, energy, etc.). In most cases it is used to represent a single dimension; ecological/environmental. “Sustainability is the study of how natural systems function, remain diverse and produce everything it needs for the ecology to remain in balance (Environmental Science, 2019).” The agreed term by United Nations (UN) for sustainability is defined in the form of Sustainable Development Goals (SDGs), to be achieved by 2030, compromising 17 goals and 169 targets.

According to (Brundtland, 1987) report, also known as “Our Common Future” for the World Commission on Environment and Development (WECD), Sustainable Development (SD) is defined as the development that meets the needs of the present without compromising the ability of future generations to meet their own needs. This definition is supported by the International Institute for Sustainable Development (IISD). SD in context improves “the quality of human life while living within the carrying capacity of supporting ecosystems” (IUCN/UNEP/WWF, 1991). As (Shaker, R. R., 2015) states, “The term “Sustainability” should be viewed as humanity’s target goal of human-ecosystem equilibrium (homeostasis), while “Sustainable Development” refers to the holistic approach and temporal processes that lead us to the end point of sustainability.” Sustainable development is comprised of three basic pillars – (1) Economic, (2) Environmental and (3) Social. In SD the three pillars are not mutually exclusive but interdependent and are a means to meet basic life necessities without damaging the natural environment. Economy is a branch of human society, which itself is wholly reliant on the natural biosphere for its sustenance. Description of the three pillars by (Munasinghe, 2016, p. 35), is as follows:

“The economy is geared mainly towards improving human welfare, primarily through increases in the consumption of goods and services. The environmental domain focuses on the protection of the integrity and resilience of ecological systems. The social domain emphasizes the enrichment of human relationships and achievement of individual and group aspirations."

Sustainability has emerged as one of the most important decision-making issues covering both qualitative and quantitative factors for organizations and governmental institutions. The New Urban Agenda (Habitat III) framework requires municipalities, local and federal governments to implement new rules and regulations for more sustainable development of urban areas while protecting the environment and ensuring prosperity and social wellbeing (NUA, 2016). In 2015, a landmark agreement (tabled as the Paris Agreement) was reached to combat climate change and mitigate the threat of climate change. The central target is to limit the global temperature rise within this century below 2° Celsius above pre-industrial levels and strive to further limit the temperature rise below 1.5° Celsius (UNFCCC, 2015). The principle theme of these frameworks is to address environmental challenges and the anthropogenic interaction with the natural environment, with the "Planet Earth" being the focal point alongside the social welfare of the populace and their prosperity.

Evaluation of sustainability is to elucidate what actions or decisions contribute to discerning sustainable development (Burgass *et al*., 2017). It is challenging to define the boundaries of sustainability due to the vagueness of the notion itself (European Commission – Joint Research Center, 2012). The three principles of sustainability constitute multiple parameters within each sphere with respect to the subject matter. In order to ensure each principle is completely represented, criteria for measuring sustainability, need to be carefully identified and drafted. Selection and application of appropriate criteria with regards to the framework of sustainability is of extreme importance and vital for the success of monitoring and regulating the effect of human actions on the environment. As stated by (Rametsteiner *et al*., 2011), "The role of sustainability indicators is to structure and communicate information about the key issues and their trends considered relevant for sustainable development." These indicators need to perfectly ally with the desired goals and objectives of the organization or community and participatory processes enable the collection and application of actors' input in the process of model building. Hence a socio-technical process is necessary to obtain all the viewpoints of the stakeholders and parties involved.

2.1.1 Sustainability Assessment

Apart from incorporating the multiple dimensions (social, environmental and economic), sustainability assessment has to inculcate cultural and value-based elements. Sustainability is still an emerging field and the dire urgency of global warming has attracted global efforts, boosting research in the field of sustainability and finding newer and better ways of meeting the world's growing energy and consumer demands while simultaneously reducing polluting emissions. (Srinivasan *et al*., 2011) says that several research efforts in the field of sustainability, particularly in environmental decision-making, performance monitoring, policy evaluation and benchmarking comparisons, are evolving within the scientific community. Sustainability of a system cannot be assessed by the use of a single criterion mainly because of its intrinsic multidimensionality characteristics thus an evaluation incorporates all three dimensions (economical, environmental and social) needs to be considered for a more comprehensive assessment of sustainability (Doualle *et al*., 2015). Policies considering sustainability involves numerous decision-makers form heterogenous backgrounds with differing values. Sustainability Assessment (SA) is conducted to support decision making process and policy development and has become increasingly common in product and institutional assessment (Sala *et al*, 2015). It is imperative for the decision support tools to be flexible enough to accommodate differences in decision making styles catering local and national interests concerning economic, social and environmental impacts.

Environmental considerations have significantly raised the demand of sustainability assessment tools and are based on a wide number of methodologies focusing primarily on the economic outcomes. However, the social and environmental principles are often neglected or handled inappropriately. They are usually listed alongside the analysis report and not evaluated which results in an incomprehensive evaluation. Assessment of products constitutes managing criteria focused on raw material consumption, production line, use of product and disposal on end of life. These are vital criteria especially concerning the PLANETIERS product portfolio.

2.2 Assessment Methods

The methods reviewed in this section are catered towards sustainability assessment while also taking into consideration the need to develop a decision support tool for PLANETIERS which encompasses all dimensions of sustainability previously addressed. The purpose of sustainability assessment is to assist decision makers in determination of actions to make the society more sustainable (Devuyst, 2001).

Life Cycle Assessment (LCA) is the most established and prevalent technique and it is comprehensive in quantifying environmental impact, however not all ecological effects are covered, only those that are allocated a functional unit (Finnveden *et al*., 2003). LCA evaluates potential impact on the environment by quantifying environmental impact and relies on measured data for calculations which can be problematic to provide. This methodology has its shortcomings as it strongly relies on quantitative data which cannot be applied to the small-scale firms and enterprises.

In Cost Benefit Analysis (CBA), the values are expressed in monetary terms and adjusted for its present-day value. In CBA approach, environmental and social factors are included but are limited to ones that can be reasonably quantified. This proves to be often problematic, similarly to LCA, due to the scarcity and difficulty in retrieving relevant data (Belfield and Levin, 2010). In practice it is hardly ever realistic to value all the costs and benefits of options in monetary terms especially concerning sustainability. Qualitative variables that are crucial for assessing sustainability are relinquished due to the fact that they cannot be monetized thus incurring substandard results in the final analysis.

Life Cycle Costing (LCC) is a method for calculating costs of process or product over its lifetime and in itself it is not affiliated with environmental costs and a modified version of LCC is needed to incorporate ecological impacts. It achieves this by translating environmental problems into a one-dimensional monetary unit. The monetarism of LCC results in loss of details which in turn limits decision-makers' comprehension of environmental problems (Gluch and Baumann, 2004). This results in incomplete and inaccurate monetary calculations with regard to environmental costs.

Cost Effectiveness Analysis (CEA) compares of outcomes of decisions in a distinct manner to CBA where it does not assign monetary values. Nevertheless, it is still susceptible to the similar limitations of CBA, as relevant data is hard to collect (Belfield and Levin, 2010). Poor quality or lack of data inhibits the formulation of the most beneficial environmental decisions.

Although these methods have been employed for analysis and impact assessment, they are not suitable for the present case study due to the multiplicity of factors and indicators concerning sustainability which are qualitative in nature and cannot be quantified. Furthermore, the methods are not adequate for small-scale enterprises, of whom which the PLANETIERS predominantly composes, as they lack the resources and infrastructure to provide large sums of quantifiable data. Thus, a more pragmatic methodology needs to be adopted that takes the aforementioned limitations into consideration.

2.2.1 Multi Criteria Decision Analysis

Keeney and Raiffa introduced Multi Criteria Decision Analysis (MCDA) techniques in 1976 as an approach towards structuring of objectives and generation of attributes (in this particular case "attributes" is referred to descriptors of performance). The final objective is to recognize the optimal decisions by identifying tradeoffs between them. (Dodgson *et al*., 2009) describes MCDA as "both an approach and a set of techniques, with the goal of providing an overall ordering of options, from the most preferred to the least preferred option." Its advantage is that it can incorporate both qualitative and quantitative in the process. (Keeney and Raiffa, 1993) states that MCDA enables structuring of objectives by differing areas of concern in order to encompass multiple points of view. (Bana e Costa and Vansnick, 2000) noted that the judgments of actors concerning the attractiveness of decision opportunities and alternatives plays a crucial role. MCDA accounts those judgements and is suitable and preferable compared to other evaluation methods on the account of its flexibility and its feasibility to engage in dialogue between stakeholders, analysts and experts. MCDA has the appropriate set of features to handle the interrelationships between indicators to assess sustainability while in consideration of different values uncertainties, perspectives and stakeholders. MCDA has a technical aspect which employs a set of technics in the evaluation process and a social component which captures the points of view of stakeholders involved (Phillips and Bana e Costa, 2007). As previously stated, sustainability assessment incorporates multidimensional characteristics with the existence of multiple criteria bringing about significant complexity in the decisions making process. In the case of numerous options, decision makers are not able to understand them comprehensively, due to the natural limitations of perception capability (Phillips and Bana e Costa, 2007). MCDA addresses this owing to its cyclical nature, enabling new understanding and comprehension, hence structuring of complex problems and leading to informed and better decisions.

Sustainability assessment has been excessively carried out by MCDA due to its flexibility and allowance for dialogue and stakeholder interaction (Cinelli *et al*., 2014). It has been widely used in assessment of electricity technology (Klein and Whalley, 2015; Roinioti and Koroneos, 2019), energy (Myllyviita *et al*., 2013; Padilla-Rivera *et al*., 2019; Martin-Gamboa *et al*., 2017), waste management (Gabbay de Souza *et al*., 2016), water treatment (Akhoundi and Nazif, 2018; An *et al*., 2017), agriculture (Irene De Luca *et al*., 2017) and construction and mining (Joglekar *et al*., 2018; Yaylaci and Duzgun, 2016). In recent years with regards to environmental and ecological concerns, MCDA is one of the most practical methodologies for decision making. MCDA methods are able to combine both qualitative and quantitative information and, in this regard, can overcome some of the limitations faced by other methodologies, which often reduce all environmental dimensions to one sole unit of analysis. In sustainability assessment, economic, environmental and social indicators are used as criteria to assess the different alternatives in an integrated manner. In addition, software and decision support systems have been developed to aid in application of MCDA techniques.

(Munda, 2005; EPA, 2006), quoted by (Cinelli *et al*., 2014), states "LCA is a product-oriented tool for the assessment of environmental implications, while multi criteria decision analysis (MCDA) is a set of methods that can be used to compare alternatives from a product level to a policy one, by covering one or more sustainability pillars." It is a well-structured and is capable of integrating opinions of a heterogenous panel of stakeholders, experts and decision makers. MCDA is preferable when the alternatives to be evaluated against a set of attributes are hard to quantify. Most importantly, MCDA enables us to accommodate qualitative indicators alongside quantitative ones, thus employing a more holistic and detailed methodology in developing a framework pertinent to the scope of this dissertation. Due to the complexity of sustainability assessment and the non-numeric nature of the social and environmental indicators, MCDA techniques have great potential to contribute in this area (Cinelli *et al*., 2014; Carayannis *et al*., 2018).

The methods found in the literature consider the alternatives under evaluation to be well-defined projects or decisions, and amongst them a singular or set of alternative need to be selected. This is not the case for the present thesis study as the alternatives are virtually infinite and it would be appropriate to construct a tool that employs a method which allows systematic consideration of all alternatives. For this reason, MCDA techniques provides a suitable method for assessing sustainability of household products and commodities.

2.2.1.1 MCDA Techniques

The techniques of MCDA mentioned below have a shared commonality, which is taking into consideration the multidimensionality of the decision problems, however, each method possesses its own properties with respect to assessing criteria, application and computation of weights, utilization of mathematical models and the skill of participants to partake in the process.

Analytical Hierarchy Process (AHP) is extensively used in organizations and has made immense contributions in the area of multicriteria decision making, however it has come under criticism with respect to derivation of priorities (Belton and Gear, 1983), stating "*there is a degree of imprecision in the specification of what factors should be taken into account ….",* and "*in some circumstances the method can give anomalous results arising from a misunderstanding of what is required ….*". The numerical scale maybe inconsistent with the decision maker's interpretation of the verbal expressions and often do not lead to normatively correct numerical results. Furthermore, biases may occur due to misinterpretation of the verbal expressions. The phenomenon of "Rank Reversal" is manifested when adding an irrelevant (Perez *et al*., 2006). (Bana e Costa and Vansnick, 2008) concludes that the Eigenvalue Method (EM) has a central weakness thus rendering the use of AHP very problematic. In the AHP approach the weight coefficient against each alternative is obtained with respect to the overarching goal at the top of the hierarchy, and as there is no unique and explicitly defined singular goal in the terms of sustainability assessment, hence AHP is not a viable approach.

Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), requires that the attributes be described numerically or should be easily transformed into calculable units (Bystrzanowska and Tobiszewski, 2018). This is not practical as the criteria for sustainability assessment are qualitative in nature.

Goal Programming (GP) has the capacity to handle large-scale problems and operates with the objective function of achieving a set of predefined goals as closely as possible. GP's major disadvantage is its inability to weight coefficients and has been used in combination with other methods to accommodate this drawback (Velasquez and Hester, 2013).

Elimination and Choice Expressing Reality (ELECTRE) and Preference Ranking Organization Method for Enrichment of Evaluations (PROMETHEE) are outranking methods is that they require a pairwise comparison of alternatives, which limits the number of alternatives to be considered. They can filter out the less favorable options and identify the leading alternatives. However, in the scope of the thesis, this feature is irrelevant as the goal of the thesis is to assess the sustainability of all current and future products. Henceforth, the outranking methods are rejected in the present case study.

Multi-Attribute Value Theory (MAVT) is used to address problems involving multiple alternatives that have to be evaluated on the basis of conflicting objectives. MAVT is able to handle quantitative and qualitative data, hence considered vital in the field of environment policy making, where many aspects are often intangible. In case of non-availability of quantitative data, expert judgments can be applied to estimate the impacts on a qualitative scale. The decision process in MAVT is able to include different perspectives and facilitate discussion when the decision context is complicated and consists of conflicting stakeholder views. The goal of MAVT is to construct a means of associating a numeric value with each alternative, in order to produce a preference order on the alternatives consistent with the judgments of the stakeholders (Ferretti *et al*., 2014). MAVT methods aim at aggregating different functions and utilize mathematical models to achieve that aim in the evaluation phase of the model building process. To realize the mathematical model, MAVT users are requested to assign weights to the attributes that reflects their relative importance, irrespective of the units or outcomes of the attributes. This procedure allows forth the compensation of a weak performance of one criterion by a good performance of another criterion, hence constituting MAVT as a compensatory technique. It is important to note that this compensatory approach is crucial in the field of sustainability assessment (Giuseppe, 1995).

Multi Attribute Utility Theory (MAUT) is considered an extension of the MAVT and accounts for risk measurement. MAUT takes into consideration the decision maker’s preferences in the form of the utility function which is defined over a set of attributes (Pohekar and Ramachandran, 2004). Uncertainty is modeled probabilistically as a distribution over the possible attribute values for a given action on the part of a decision maker. The determination of consequences in MAUT can be a very time consuming and difficult, sometimes an impossible task especially if it consists of an exceedingly high number of attributes because of the problems incurred in specifying the appropriate set of attributes and measurement procedure, in addition to the complex, obscure and uncertain relationships that often exist between decisions and their consequences (Delforce and Hardaker, 1985). This approach is suitable for accommodating uncertain events and consequences. However, this desirable feature does not pertain to the context of this dissertation as the impact of the commodities on social, economic and environmental fronts has already been established and predetermined.

Measuring Attractiveness by a Category-Based Evaluation Technique (MACBETH) can accommodate quantitative and qualitative data amidst a heterogenous set of criteria and involvement of multiple stakeholders. The MACBETH approach requires pairwise comparisons on an interval scale with a strict consistency check and then uses linear optimization to calculate the priorities (Ishizaka and Siraj, 2018). If the linear program is unfeasible, denoting that the pairwise comparisons are inconsistent (Ishizaka and Siraj, 2018). The priorities cannot be calculated at all when the DM is inconsistent and needs to be resolved beforehand to proceed onwards. The systematic inconsistency checks prove to be very helpful as they offer a great opportunity to reconsider one’s judgements. The use of numerical depictions of preferences can either be too complex for experienced facilitators and actors may also find it more intuitive and less time consuming to provide qualitative responses (Bana e Costa and Chagas, 2004). Therefore, less experienced actors may derive significant benefit from the use of a “client friendly” approach as a path to cardinal value measurement. MACBETH approach has been used for sustainability assessment purposes (Dhouib, 2014; Marques *et al*., 2015).

Problems comprising of a solution from a finite and determined set of options, usually employs outranking or MAVT methods. The MACBETH approach is selected to be the most adequate and viable method for the development of the multicriteria model in consideration of the relevance of the techniques pertinent to the scope of this thesis and it is expanded in great detail in the segment hereinafter.

2.2.1.1.1 Measuring Attractiveness by a Categorical Based Evaluation Technique

MACBETH was authored by Carlos António Bana e Costa, Jean-Marie De Corte and Jean-Claude Vansnick, and permits the evaluation of options against multiple criteria. The MACBETH method has been used in various fields for decision-making and is designed to be user-friendly for group decisions (Lavoie *et al*., 2016). This method has been used successfully in analysis of many medical, municipal and engineering policy projects (Bana e Costa *et al*., 2012; Bana e Costa and Oliveira, 2002; Baba e Costa *et al*., 2008). (Bana e Costa *et al*., 2003) frames MACBETH approach as an interactive evaluation of problem and elaboration of recommendations to choose options in an individual or group based decision-making process. (Bana e Costa and Chagas, 2004) states "MACBETH employs a non-numerical interactive questioning procedure that compares two stimuli at a time, requesting only a qualitative judgment about their difference of attractiveness." The MACBETH model abstains stakeholders from having to directly assign numerical scores, found prevalent in other techniques. This characteristic is of particular value when incommensurable objectives of public decision analysis involving economic efficiency, social equity and environmental quality need to be assessed. MACBETH requires qualitative judgments about differences in attractiveness (Sanchez-Lopez *et al*., 2012). The MACBETH approach proposes a simple interactive inquisitive process to elicit quantification of values for the construction of a value scale through pairwise verbal judgements of difference of attractiveness between elements. The evaluation executed in terms of the user perspective is in the form of an interval scale. (Bana e Costa *et al*., 2003) mentions that the qualitative judgements about difference in value helps to quantify the attractiveness of options. The decision-maker's judgments are subsequently represented as a numerical scale and through a similar process it allows forth the creation of weighting scales for criteria (Bana e Costa and Chagas, 2004). (Lavoie *et al*., 2016) states that the MACBETH approach is especially helpful in achieving consensus among various stakeholders and is an extremely effective channel to accurately model the ideas expressed by a group of experts. (Bana e Costa and Chagas, 2004) concludes that by acquiring qualitative answers, rather than quantitative ones, the technical burden placed upon the decision maker was eased and which in return appears to have had a positive impact on the level of confidence on the provided answers. (Bana e Costa and Chagas, 2004) further adds, the relative ambiguity of the possible answers was crucial as the stakeholder was forced to think carefully and prudently before making their judgement.

The MACBETH approach in the decision-making scope is presented in this dissertation as a means to develop a model to assess sustainability of everyday consumables and household commodities available on PLANETIERS' online marketplace.

2.3 Participatory Methods

(Marttunen *et al*., 2013) suggests that the more MCDA is used in an integrated and interactive way, the more likely its potential benefits are to be achieved. Participatory methods have established themselves as the prevailing approach in aiding social component of MCDA's technique, incorporating actors' input while enabling creation of a shared understanding of the issue present. Participatory processes that combine and communicate the judgements and expertise of the members in MCDA processes have greatly increased in demand. Participatory methods involve stakeholders and assist in integrating their knowledge, values, opinions and ideas. (IES, 2019) states that participatory methods may also be used to explore the knowledge base (identifying knowledge gaps), to assure relevance of the assessment, to increase the social robustness of an assessment and validate the quality of the assessment from a societal point of view in terms of purpose, relevance and legitimacy. (Hage *et al*., 2010) comments participatory approaches in environmental knowledge production are commonly propagated for their potential to enhance legitimacy and quality of decision-making processes, especially under conditions of uncertainty. Involvement of external stakeholders and experts helps assimilate different perspectives, values and knowledge base to improves the development process and a variety of participants aids in overcoming bias and explore different pathways (Ernst *et al*, 2018). These methods enhance group facilitation and are designed to explore the level of consensus among members. MCDA methodologies have further facilitated in creating conditions for meaningful and effective interaction, which has been found to be one of the key objectives for designing participation processes (Marttunen *et al*., 2013). (Ernst et al, 2018) mentions that participatory methods can enrich the process by bringing clarity to the various and contradictory values, and underlying problems of the participants involved. According to (Rowe and Frewer, 2000) participatory engagement needs to respect key normative criteria, including representativeness and transparency to underpin legitimacy.

All participatory approaches of this social processes, possess both iterative and interactive characteristics, complementing the technical components of a MCDA methodology. The choice of participatory method depends a great deal on the defined objective, goals and issues to be addressed, problem context and participants involved. Participation processes should be viewed with two main perspectives; exploratory, which is investigative in nature with no connection in policy-making and deliberative, which is utilized for decision-making. Each participatory method requires a specific style of moderation that will affect how the process is conducted and the resulting outcomes (IES, 2019). Exercising different participatory social approaches to collect information may consequently result in the generation of distinct outcomes.

In decision conferencing, the facilitator moderates the whole process but does not actively partake in the process or portrays their opinion while the actors of the process contribute the content. A decision conference is led by a neutral facilitator, whose role is to enhance communication between the participants and to get them to constructively deal with the conflicting issues at hand (Marttunen *et al*., 2013). In decision conferencing, it is presumed that the actors possess all the necessary information to resolve the problem at hand. These conferences typically last two or three working days, but can also be concluded in a single day. (Phillips and Bana e Costa, 2007) remarks that the model represents the collective view of the group at any point during its generation and modification and as it is projected for all participants to see as it is created, it helps gain confidence in model results. Shared understanding among the actors is generated as the model building process advances. The aim is to combine group process facilitation, preference modeling and information technology with the concept that a group can achieve better results than actors working individually (Marttunen *et al*., 2013). Decision Conferencing has been used repeatedly with MCDA in (Elghali *et al*., 2007; Barfod and Salling, 2015; Reddy *et al*., 2014) among others. The use of decision conferencing with the MACBETH approach has been found in (Rodrigues, 2014).

In decision interview it is suggested that this is carried out in close co-operation with the stakeholders. The open discussion with the stakeholders and experts on the impacts can help to crystallize the reasoning behind them and pinpoint gaps in knowledge and previously ignored uncertainties (Marttunen *et al*., 2013). This approach tends to gather a significantly large amount of information and drawing out recommendations and conclusions is a delicate task. During the interview it is vital that the facilitator inquires and records the arguments of the interviewees with an aim to understand the reasoning behind the importance of the attributes and consequently the desirability of the alternatives (Marttunen *et al*., 2013). It is suggested that the groups of similar opinions are identified by presenting the results at an individual level. This method facilitates the understanding of all stakeholder viewpoints and helps illustrate the discrepancy of the preferences. Decision interviews are usually part of an iterative participation and learning process aimed at building common understanding and finding or constructing broadly acceptable alternatives through the identification of key trade-offs and balancing between important objectives (Marttunen *et al*., 2013). The lack of standardized approach makes aggregation of data difficult and less precise. It is practical when the number of experts is small.

The Delphi method ensures anonymity during actors' involvement in the model building process, thus preventing conformity and external influence in the group's input. A facilitator organizes a conference, compiles reactions and shares results or comments with the group. It is the aggregation of opinions from diverse fields and departments to come to a common true goal. The main objective should be to select panelists with the capability, knowledge, professional qualifications and relevant experience in the field under investigation. The facilitator moderates the whole process but does not take part in the active decision process. Its iterative nature makes it extremely resource and time consuming. This approach is expanded in great detail in the segment hereinafter. Delphi approach has been extensively used in numerous fields alongside MCDA (Van Schoubroeck *et al*., 2019; Doukas *et al*., 2006; Ivlev *et al*., 2015). The Delphi method has also been utilized with the MACBETH approach in (Vieira *et al*., 2019).

The Nominal Group Technique (NGT) is a structured "*face to face*" technique, where a the research team formulates a nominal question and records the participants responses privately, who are then requested to share their answers leading to a group discussion and subsequently a voting round, whose results are presented to the participants anonymously (Foth *et al*., 2016). NGT is a qualitative technique, used to generate ideas and rank the items. The feedback mechanism in NGT has a manipulating effect towards participants leading to conformity instead of consensus.

Focus Group Technique (FGT) engages small groups of people to discuss a certain topic under guidance of a moderator. Expert opinions are elicited and post-group questionnaires are distributed to gather overall opinions of the participants. Focus group meeting are highly accountable and the members are requested to justify their points of view. It is possible that members may try to minimize conflict and may fail to share their knowledge without critically analyzing and assessing the information (Baron, 2005). According to (Gardner *et al*., 1999), there is the potential of focus group's views being influenced by the group dynamics.

Participatory methods designed for group interactions can be conducted either through being physically present (e.g. decision conferences, decision interviews and Delphi) or non-presential (e.g. Delphi) and should be chosen when the objective is to collect judgments on predetermined issues. As mentioned under sustainability, selection and application of appropriate criteria is necessary for the formulation of the multicriteria model, henceforth, expert opinion and knowledge needs to be elicited from specialists and professionals. The Delphi process is the most adequate form of participatory method to achieve this goal. The MACBETH approach requires verbal judgements emanating from stakeholder's preferences which requires long hours of discussion and dialogue. To facilitate this socio-technic approach, decision conference is the most feasible option.

2.3.1 Delphi

Developed in the 1950's, by Olaf Helmer, Norman Dalkey and Nicholas Rescher of the RAND Corporation (Santa Monica, California) under the US Army Air Corps, to predict the future technological capabilities of the military forces. RAND researchers developed a structured survey in written form in order to estimate bombing requirements. The first evaluation from the panel of experts did not result in a consensus, however, consensus was achieved in the second evaluation and the procedure was said to have yielded more reliable results than comparable techniques present (Von der Gracht, 2012). The procedure was declassified in the sixties by the American Armed Forces where it was previously reserved for military use, resulting in its rapid growth and widespread use, both geographically and thematically, particularly towards evaluation of complex social problems (Landeta, 2006). The name is derived from the “Oracle of Delphi”, a priestess in the temple of Apollo known for her prophecies in ancient Greece.

Its initial objective was to obtain the most reliable consensus among a group of experts. As exclaimed by (Dalkey and Helmer, 1963), it achieved this through “… a series of intensive questionnaires interspersed with controlled opinion feedback.” It is a process of which can help in arriving at a group consensus by providing questionnaires and subsequent responses. (Gnatzy et al., 2011) states that the conventional Delphi can be defined as a method that aims at a consensus on a particular topic among a group of experts. It is based on the principle that forecasts or decisions from a structured group of individuals are more accurate than those from unstructured groups (Rowe and Wright, 2001). The widespread diffusion of Delphi process has resulted in consensus measurement no longer being the technique's primary aim, yet the consensus measurement still has to be considered an important component of Delphi analysis and data interpretation (Von der Gracht, 2012). In (Linstone and Turoff, 1975), Delphi is mentioned as a method for structuring a group communication process and enabling group of individuals to deal with a complex problem effectively, whilst (Reid, 1988) states the Delphi is a process for systematic collection of informed judgements from a group of experts on specific problems As the usage and the number of modifications to the approach have increased, there are now many different versions in existence and not all Delphi techniques aspire to achieve consensus, conversely it may attain dissensus. (Linstone and Turoff, 2010, p. 1714) states “In fact, a bipolar distribution may be a result and a very significant one indeed”. The aim of Delphi process in (Coates, 1975), quoted by (Linstone and Turoff, 2010) is as follows:

“The value of the Delphi is not in reporting high reliability consensus data, but rather in alerting the participants to the complexity of issues, by forcing, cajoling, urging, luring them to think, by having them challenge their assumptions…... [M]ore attention should go into the basis of divergence rather than the basis of convergence.”

Delphi enables input from a larger number of participants that can either be present in a form of a group or board meeting, allowing the respondents to answer concurrently or from those members who are geographically dispersed to engage in the questioning process in an asynchronous way, by enabling them to answer on their own accord depending on their time availability.

The Delphi procedure is defined by four basic principles: 1) Anonymity, 2) Iteration, 3) Controlled Feedback of responses to all group members, and 4) Statistical Aggregation of Individuals' Responses (Rowe and Wright, 2001). The undermentioned attributes are key characteristics of the Delphi procedure, leading towards an effective decision-making process.

- The responses are kept anonymous allowing the members to convey their opinions candidly while simultaneously discarding any form of undue pressure. The key feature of the Delphi process is its aspect of anonymity, allowing the actors to freely express themselves hence distancing the domineering effect present in the hierarchical structure of employees within the corporate sector. Anonymity in surveys usually leads to higher response rates (Von der Gracht, 2012).

- Iterative questionnaire sessions over a number of rounds, allow for adjustments to the original answer. Regarding the optimal number of iterations, (Belton *et al*., 2019) recommends three rounds should generally be sufficient to allow a pattern of stability to emerge from panelists. The respondents are provided with an opportunity to either maintain or change their previous stated responses in the wake of group feedback.

- The feedback constitutes the judgements of all the group members and informs them the opinions of their fellow participants. The feedback is termed “controlled” because the facilitator decides on the type of feedback and its provision (Von der Gracht, 2012). The Delphi process enables the facilitator to control the flow of information and filtering out the irrelevant content hence avoiding the negative side of group dynamics. In case when decisions strongly deviate from the group response, participants usually provide reasons for their unique perspective.

- (Von der Gracht, 2012) mentions that “the statistical group response can be presented either numerically or graphically, and usually comprises measures of central tendency (median, mean), dispersion (interquartile range, standard deviation), and frequency distributions (histograms and frequency polygons).” Analysis of the data over successive rounds allows for measuring not only the existence of consensus and its strength, but also the convergence of opinions.

Delphi has undergone numerous developments since its conception. The extensive employment of Delphi across a broad spectrum of areas has influenced a series of changes in order to adapt to differing scenarios (Linstone and Turoff, 2010). The characteristics of each Delphi design can be referred from (Hasson and Keeney, 2011). Sustainability is an extremely broad field constituting of multiple aspects within its domain, hence it is necessary to attain specialist opinion from a diverse group of experts. It is neither feasible or practical to consult the experts individually or person. For these reasons, “Web-Based Delphi” design has been identified to be implemented to develop the multicriteria model. Web-based Delphi design is a very practical approach as it can incorporate a very large number of participants and also geographically dispersed members. This is suitable for the facilitator and the respondent as they can respond in their own convenience. The Delphi multi-round survey procedure has been widely and successfully used to aggregate expert opinions and will be used in the scope of this thesis to gather feedback on the relevance of criteria with regards to evaluating sustainability of household items and consumer products.

3 Framework Methodology

In the previous section, various methodologies were mentioned, however no literature was present regarding sustainability assessment of products, hence a socio-technical methodology needed to be developed for building of the tool. A new methodological framework is proposed for building a sustainability assessment tool in order to appraise products. PLANETIERS will be the case study for testing the methodology and the tool. This model would allow for ranking of products based thus enabling approval and selection of sustainable products (and rejection of non-sustainable merchandise). The sustainability assessment framework combines 1) technical elements of a multicriteria model built with the MACBETH which is built upon the principles of multi-criteria value measurement, with 2) social elements of participatory processes for capturing expert opinion on relevance of criteria for appraising products. The framework requires identification and establishment of criteria which will serve as input parameters to the decision-making model. This section includes the methodology undertaken for solving of the decision problem at hand.

3.1 Outline

As mentioned, little or no literature was available regards to the development of the tool for assessing sustainability of products, hence bringing about complexity in this novel concept. For this reason, a framework was developed in order to approach such an intricate task by splitting the overall problem into more simple parts. It was essential to structure the framework in a streamline fashion ensuring the tasks were coherent with the global context. The main activities of the multicriteria framework in developing the sustainability assessment tool are depicted in Figure 1. The first step of the proposed framework concerns with the structuring the problem context. The next task is to gather all relevant information to identify criteria under the pillars of sustainability with relevance to the problem context. These criteria will serve as the basis for the creation of the value tree and the establishment of the evaluating criteria. Descriptors of performance need to be associated with the criteria to make them operational for evaluating alternatives. In the subsequent steps, value scales will be generated for each criteria and weights will be assigned to the criteria. The value tree being a hierarchical structure of two levels, weightages will be allotted to both levels and thus conclude the evaluation phase of the methodology. Finally, the options are appraised, leading to the analysis of the results and the selection of sustainable merchandise. Although the framework is separated into its main constituting parts, these are closely interrelated. The framework is iterative and adaptive in nature facilitating multiple points of view and hold the capability to introduce changes and adjustments where deemed necessary, especially during the structuring and evaluation stages of the model building process.

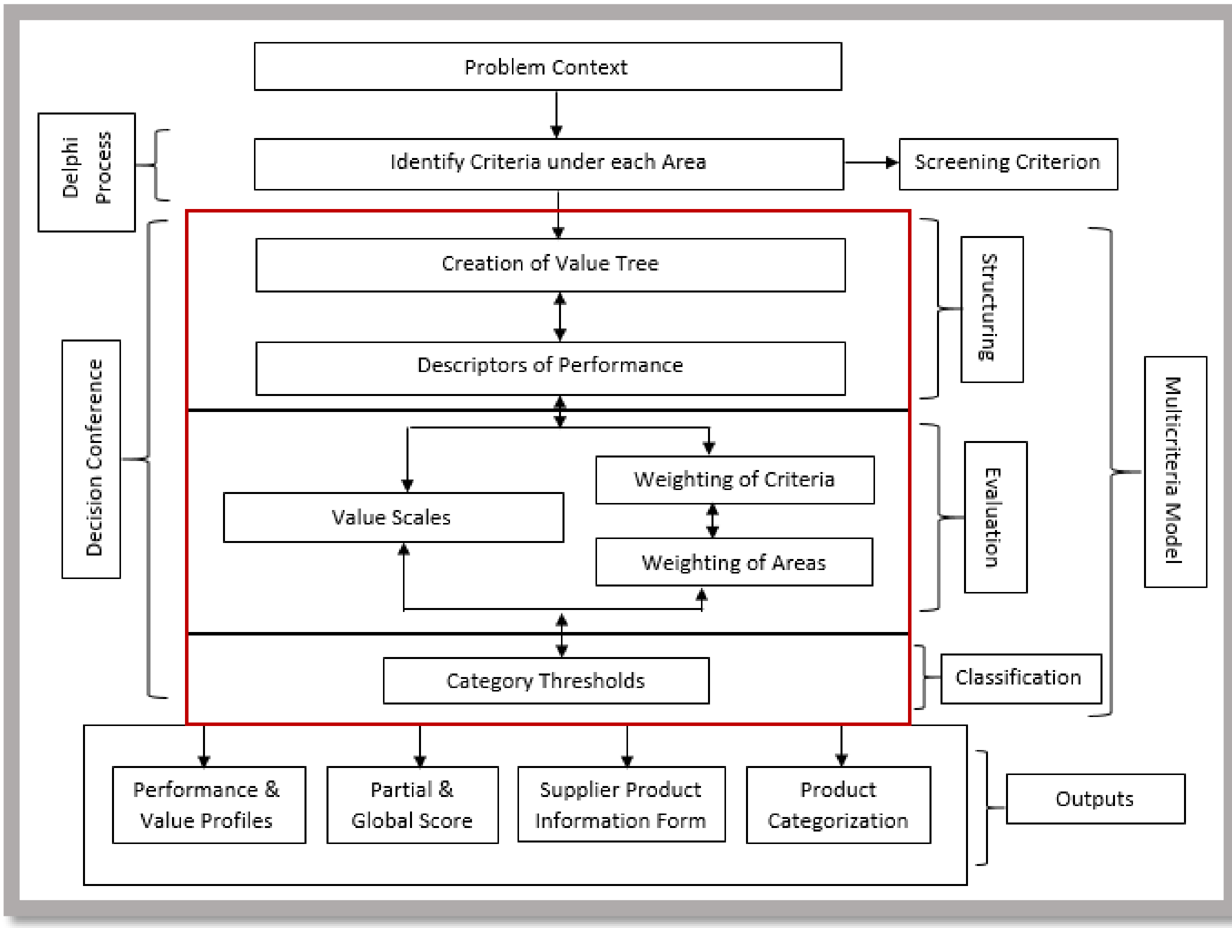


Figure 1 - Framework for Development of the Assessment Tool

The aforementioned set of activities within the multicriteria framework will result in the construction of a product assessment tool. This tool will capacitate the decision makers to discern if a product meets a minimum score to be rendered sustainable. The ensuing results of the evaluated products will be analyzed leading to the classification and categorization of products on the basis of the scores achieved.

The schematic for the operations taking place for product appraisal is displayed in Figure 2. The schematic represents a process flow diagram from moment the supplier registers their product on the upgraded online platform to the moment the decision makers analyze the results. The supplier will input the product specifications on an online form, constructed by the decision makers, corresponding to the criteria descriptors. This form will feed information to the tool which in turn will generate the results for the products to be evaluated. The stakeholders would analyze the results of the model and deicide to approve or reject the products. In this case study, products from the PLANETIERS inventory will be assessed to test the model and ascertain if the model can be considered "requisite" (Phillips, 1984). The validity of the model is attained from conducting sensitivity and robustness analysis.

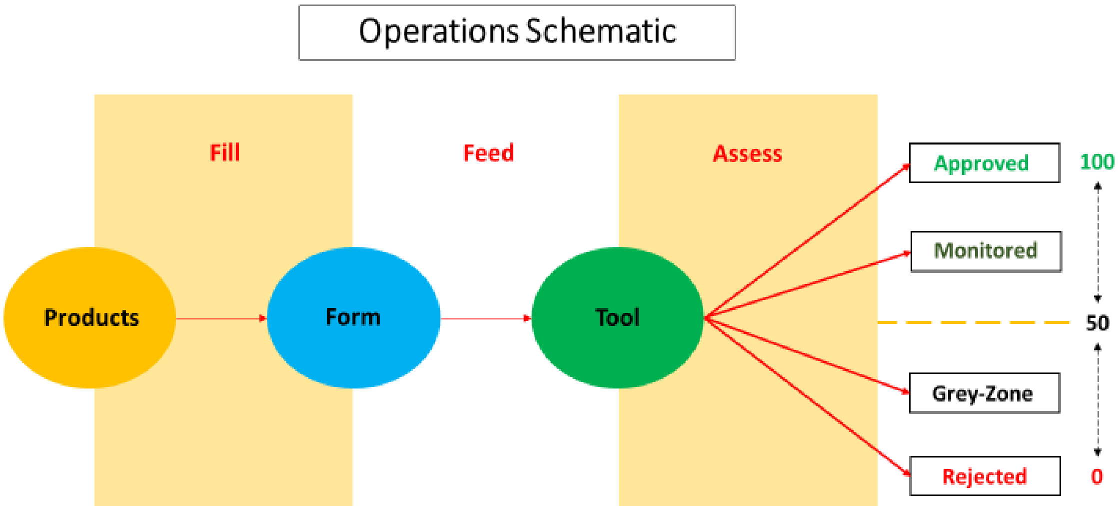


Figure 2 - Operations Schematic

On the technical side, MACBETH approach was used to design the model within a hierarchical structure, whereas the social side of the application was comprised of participatory methods incorporating web-Delphi processes and decision conferences. In the socio-technical process of the Delphi process and the decision conference, the role of the facilitator will be enacted by the thesis author.

3.2 Problem Context

Defining problems allows for measures to be undertaken in order to rectify them. The first task was to define the problem as explicitly as possible so all parties are well informed and accustomed to the main objective. The problem pertains to assessing sustainability of products. However, no tool has been established to carry out this activity, therefore, requiring construction of such a novel tool. The tool is to be based on the term "sustainability", which is characterized by the three principles and can be measured through multiple criteria, hence justifying the use of MCDA framework to properly accommodate each criterion. The conveyance of this decision problem and the objective was ensured while conducting the Delphi process, as expert and specialist opinion was elicited, and in steering the decision conference, as multiple perspectives of the stakeholders was to be accommodated.

3.3 Identifying Criteria

It is necessary to gather relevant information keeping in consideration the aspects of sustainability pertinent to appraisal of consumer products. The accumulated information will facilitate in identification of criteria which in turn will serve as basis for the evaluation criteria and the construction of descriptors of performance. The first stage of the model construction focused on identifying set of criteria and allocating them according to areas of concern. The recognition, measurement and employment of appropriate indicators concerning sustainability is of among the major challenges facing policy-makers, scientists, researchers, bureaucrats and citizens (McCool and Stankey, 2004). Indicator development is a critical step for sustainability assessment. Criteria can be adopted by two approaches: 1) Top Down approach, where a prepared set of criteria are introduced and imposed on the stakeholders, on the condition they are scientifically valid, on the contrary, 2) Participatory (Bottom Up) approach, is the identification of relevant criteria that are agreed upon by the stakeholders. Additionally, the indicators must adhere to conceptualization and operationalization, where conceptualization refers to "What" is to be measured and operationalization refers to "How" to measure and interpret it (Opon and Henry, 2019). Potential criteria were generated by conducting a detailed literature review. Furthermore, products were inspected and criteria were ascertained with regards to specification and category. In the context of this dissertation, participatory approach has been implemented to elicit opinion of experts regarding relevance of the criteria for assessing sustainability, which is further expanded in Delphi methodology in the subsequent segment.

3.4 Web Delphi

Exercising Delphi methodology was most preferable as it allowed for consolidation of a wide range of opinions from geographically dispersed individuals with diverse backgrounds whilst ensuring anonymity. Anonymity ensures there is equal freedom of expression for everyone and there are no repercussions for their choices. Anonymity avoids the drawback of members acquiescing for the sake of agreement (Goodman, 1987). Delphi process comprises of two bodies; the facilitating, which sets up and implements the required procedures, and the panel of participants, who give their input regarding the decision problem. A web-based Delphi process was initiated to ascertain the views of experts and stakeholders on the criteria considered relevant to appraise sustainability of PLANETIERS' inventory.

The first step for structuring the multicriteria model was to launch a questionnaire to a panel of stakeholders and experts, requesting them to indicate their level of agreement and disagreement, with the 34 preestablished criteria with relevance to evaluating the sustainability of a product on the PLANETIERS online platform. Henceforth, enabling the creation of a value tree and classification of the individual criterion inside the defined areas, constituting of the three basic pillars of sustainability; environmental, social and economic.

3.4.1 Panel

In this stage, the panel of experts is identified and appointed. Delphi respondents consist of groups of experts who have been selected on a criterion sampling basis or have been chosen for their expertise in a particular field. The members of the panel should be preferred on their willingness and ability to participate in the process. Experts' degree of interest in answering the questions in the Delphi survey predicts both initial response rate and subsequent dropout rate (Hasson *et al*., 2000).

The panel defined for this study constituted of the stakeholders and experts with applicable knowledge in sustainability in a variety of domains from consultants to engineers and professors to ex-politicians. Such a diverse panel provides a much more realistic response by consolidating a broad spectrum of inputs. The involvement of a heterogeneous panel of a diverse geographical setting was deemed necessary for ascertain expertise on multiple spheres of sustainability. The number of panelists was determined on the availability and their relevance to the scope of the thesis.

3.5 Structuring Phase

MCDA has been found to be a resourceful approach to examine uncertain and multiple conflicting objectives (Watrobski *et al*., 2019). Hence, given the problematic of the case study we are inserted in, the use of a multicriteria approach was a natural and convenient course of action. The perspective of using of using a multicriteria evaluation model that will enable not only the assessment of the performances of the various considered criteria, but also its partial and overall value was introduced. A decision conference was conducted in order to validate the outcomes of the Delphi process and construct a multicriteria evaluation model. This served the purpose of assigning value scales to different performances within a single criterion as well as weighing of criteria and areas. The actors of the process collaboratively construct the criteria, performance and value scale, and assign weightage to the criteria. It should be noted that this intense level of communication necessitates the need for an experienced and relevant panel. This interactive and recursive process involving the decision makers fosters ownership of the model. This combines the technical elements of the approach with the social aspects of the interaction between the actors. This socio-technical process enhances communication within the organization and generates a sense of common purpose and shared accountability.

3.5.1 Value Tree

The structuring of a multicriteria decision-aid model entails several procedures and consists of an interactive and learning process, involving multiple criteria, representing the objective components of the problem and the subjective notions of the actors through which the facilitator helps the decision-makers in solving the problem context at hand. The established criteria should encompass several properties; consensual, where all actors should agree with the criteria; decomposable, it is mutually preferential independent and each criterion should be operational and measurable; exhaustive, it should be complete and all relevant aspects should be defined; and non-redundant, it should be concise and, features previously evaluated in other criteria should not be repeated and removed (Bana e Costa *et al*., 1999). A criterion is considered complete if it is adequate in indicating the degree to which the overall objective is met (Keeney and Raiffa, 1993). The operationality of the criterion is defined by the ability of the criteria to be meaningful to the decision maker and is able to comprehend the implications of the alternatives (Keeney and Raiffa, 1993). Redundant criteria should not be present in the model as it leads to double counting of the consequences. Redundancy can occur when "means-ends" relationship is not clearly defined and criteria are included associated with both means and ends objectives (Keeney and Raiffa, 1993). Criteria should capture the objectives and concerns of the actors and possible consequences of chosen decisions. The facilitator has to ensure all criteria are well defined and clearly apprehended by all members so as to avoid misinterpretation. The facilitator is also tasked with analyzing their relevance during the model structuring phase.

3.5.2 Descriptors

To define the structure of the criteria, key issues and concerns are identified with the decision maker. These concerns are analyzed and descriptors are developed to evaluate the alternatives (Bana e Costa and Beinat, 2005). Descriptors of performance were associated with the criteria to make them operational for sustainability assessment. Plausible performance levels are generated for each criterion to make them operational for evaluating alternatives. The levels of descriptor should be rank ordered in terms of their appeal with a decreasing order of attractiveness. In order to appraise alternatives, a level of descriptor from each criterion is assigned to an alternative.

The objective of the problem often determines the appointment of descriptors to the criteria. In order to operationalize a criterion, it is associated with a descriptor of performance. Carefully defining suitable impact descriptors is, in itself, a vital activity and several descriptors can be applicable to a single criterion, hence it is critical to select only one to avoid redundancy problems (Bana e Costa *et al*., 1999). A descriptor is a set of levels that must be ordered in terms of their partial attractiveness. A descriptor is an ordered set of plausible impact levels associated with a criterion, intended to operationalize the criterion by enabling the appraisal of options, either qualitatively or quantitatively, for which the criterion is satisfied (Bana e Costa and Beinat, 2005). Also serving the purpose to describe the impact of the option in thorough detail and in an explicit manner, so that it is unambiguous and can be perceived easily. (Bana e Costa and Beinat, 2005) states that the descriptors should ensure the ordinal independence of the corresponding key-concerns and restrict the impact levels from most desirable to least desirable thus framing the evaluation model.

A descriptor usually constitutes of three types of characteristics. First type is that it can be "*direct*", it directly reflects the effects or is related to a criterion in a natural way, or "*indirect*", indicates causes more than effects, or an index relating several indicators, or "*constructed*", describes underlying characteristics of the criterion in the form of reference levels defined by combination of several states of criteria (Bana e Costa *et al*., 1999). Second type of characteristic it classified as is "*qualitative*", utilizes semantic expressions, or "*quantitative*", uses numbers, or "*pictorial*", employs visual representations (Bana e Costa *et al*., 1999). Finally, it can be categorized under "*continuous*", represented by a continuous function, or "*discrete*", represented by a finite set of impact levels. For quantitative descriptors, finite range is defined for the intervals for comparative judgments. In the case when the performance associated with an alternative is out of the range, it defines the value score through linear extrapolation. Performance scales were developed by combining different types of descriptors which enabled criteria operationalization.

(Bana e Costa et al., 1997) suggests using two particular reference levels when generating descriptors; Good, considered attractive by the actors, and Neutral, that is neither attractive nor repulsive. The two reference levels allow comparison of attractiveness between levels. Two reference levels were then defined for the descriptors – the ''Good'' and the ''Neutral'' levels – having functional meanings for the assessment tool, to permit the appraisal of the options on each criterion and on each area. The ''Good'' reference level, envisaged by the stakeholders, indicated highly desirable properties in that particular criterion but allows for much more desirable properties. The ''Neutral'' level of performance neutral is judged to be adequate and acceptable in terms of the product assessment and is not considered deficient or substandard. The use of "neutral" and "good" reference levels in each criterion offers another advantage of developing a qualitative swing-weighting process based upon fixed references (Bana e Costa and Chagas, 2004).

3.5.3 Evaluation Phase

MACBETH approach is utilized in this phase to build value functions using verbal semantic judgements elicited from the stakeholder. Criteria are assigned weights initially at an individual level, then at a criteria cluster level and lastly at an area level.

3.5.3.1 Building Value Functions

The partial score (or attractiveness) of the options in each criterion has to be measured. To achieve this, cardinal value functions is constructed with respect to the descriptors of impact previously outlined. MACBETH appraises the performance levels through qualitative pairwise comparison questioning process by requiring the participants to indicate the difference in attractiveness in the previously specified criteria. Value functions represent scales of value scores which require quantitative depiction, assigned to level of performances, to analyze the attractiveness of the alternatives. Cardinal value function is determined by filling the judgment matrix for each criterion, as presented in Figure 3. The verbal mode that MACBETH offers is based on seven semantic categories of difference. The appraiser is asked to verbally judge the difference in attractiveness between two levels, "*x*" and "*y*", such that "*x*" is preferable or equal to in attractiveness in comparison to "*y*" (Bana e Costa and Oliveira, 2002). The panel has the option to respond from the six semantic judgements; "*Very Weak*", "*Weak*", "*Moderate*", "*Strong*", "*Very Strong*" and "*Extreme*". If the panel perceives that there is no difference of attractiveness between two levels of performance, then a judgement of "*No*" difference is also available. It is preferable to fill in the whole matrix, the M-MACBETH software allows the capability of only filling the diagonal and the last column and by using an algorithm is able to complete the judgement matrix by transitivity. Judgmental disagreement or hesitation between two or more consecutive categories is also allowed (Bana e Costa *et al*., 2012). In case when decision-maker is undecided in creation of value function when choosing the range of qualitative judgments, the MACBETH approach allows for an additional feature to select the "more than one" semantic judgements to showcase hesitation. In example, combination of "strong/very strong", which should be interpreted as "strong or very strong" rather than "strong to very strong". The value scale for each criterion is represented in an interval scale.

	L1	L2 = Good	L3	L4 = Neutral	L5
L1	no	moderate	strg-vstr	v. strong	vstrg-extr
L2 = Good		no	↓ strg-vstr	↑ moderate	v. strong
L3			no	moderate	v. strong
L4 = Neutral				no	v. strong
L5					no

Figure 3 - Judgement Matrix (Adapted from Bana e Costa et al, 2008, "Development of Reusable Bid Evaluation Models for the Portuguese Electric Transmission Company")

After filling the judgment matrix, the M-MACBETH software is able to compute a quantitative value function based on the qualitative judgments corresponding to an interval scale. The scores assigned to the different levels of performance are to be interpreted as the magnitude of difference between options. The value scale levels were dragged and adjusted, while maintaining the compatibility with the judgement matrix, until they were agreed and accepted unanimously by the decision-makers. The MACBETH numerical scale anchored on the two predefined reference levels, assigns the scores 0 and 100 to the lower and upper references, respectively. (Bana e Costa and Chagas, 2004, pp. 323) adds "As the answers are given, their consistency is verified, and a numerical scale that is representative of the decision maker's judgments is subsequently generated and discussed."

3.5.4 Criteria Weighting Procedure

The determination of weights is a key task in the construction of the model, as this allows forth the evaluation of options by aggregation of their local scores into an overall score. There are three main techniques for weighting criteria; trade-off procedure, swing weighting and the MACBETH method. In the scope of the thesis only MACBETH method was employed and thus expanded upon.

In order to assess the product at the level of area of concern, the value scores on the criteria of the area need to be aggregated. This requires harmonization of the value scores to a common value measurement scale. This is managed through application of scaling constants or weights assigned to the criteria, which when multiplied by the value scores and totaled, define the additive aggregate model. To establish consonance, scaling constants are assigned to partial values of each option across all criteria. The calculation of weighting coefficients (scaling constants) for evaluation criteria enables the construction of the evaluation model and permits the aggregation of the local scores of an option into an overall score. Similar approach to the creation of the value scale is again employed to determine the weighting coefficients. The facilitator elicits judgements from the decision maker to fill the judgement matrix by pairwise comparison.

The structure of the value tree and the number of evaluation criteria can influence the weighting scheme utilized. The decision maker is first requested to order the criteria in terms of their preference levels. The actors are told to assume that each criterion is at a "Neutral" level and then requested to select which criterion would they prefer most to improve from "Neutral" to "Good" performance level. They will then be inquired to compare consecutive criteria and share the degree of importance of improving from "Neutral" to "Good" of performance level on each consecutive criterion. The panel can choose from the seven semantic judgements. The judgement matrix has an additional column at the end, labelled "Neutral". This column denotes that each criterion is at its "Neutral" performance level. The appraiser is then asked the importance of improving the performance of that criterion. The seven semantic categories are available to choose from. Figure 4 showcases an example of weighting judgement matrix. In the M-MACBETH software, after the matrix has been filled, an interval scale is generated. The criteria on the interval scale, apart from the "Neutral" and the most preferred criteria, can be dragged and adjusted, similarly to creation of value scales, while maintaining the compatibility with the judgement matrix. The M-MACBETH software generates the weight of the criteria through normalization of the scores on the interval scale.

	[Suitability]	[Execution]	[Price]	[Neutral]
[Suitability]	no	strong	strong	v. strong
[Execution]		no	very weak	moderate
[Price]			no	moderate
[Neutral]				no

Figure 4 - Judgement Matrix (Adapted from Bana e Costa et al, 2012, "MACBETH")

The facilitator should then validate the generated weights with the decision maker, introducing adjustments if deemed necessary. The judgments elicited are anchored to the "Neutral" and "Good" references defined in each evaluation criterion which is a fundamental aspect to guarantee the validity of the process. It is paramount that the facilitator inquires about the importance of improving the swing of criteria, rather than the importance of criteria which would be misleading.

3.5.5 Additive Aggregate Model

(Bana e Costa *et al*., 2003) states MACBETH adopts the additive value aggregation model as its technical parameters are easily understood and it allows for handling of an intricate problem constituting of critical criteria in a precise and perceptible manner. (Bana e Costa and Beinat, 2005) specifies that "additive aggregation requires that each criterion be an isolate evaluation axis", that is it can hold mutual cardinal independence as ordinal independence is not potent enough. Quintessentially, the additive model consists of allocating a certain option, a partial score in each of the evaluation criteria, which are in turn multiplied by their respective weighting coefficients, yielding a global aggregate score for that option. The mathematical model enables the assessment of alternatives in obtaining an overall score. The most common mathematical formula denoted for additive aggregation model is as following:

$$v(x) = \sum_{i=1}^{m} \sum_{j=1}^{n} k_i \times w_j \times v_j(x_j) \ with \begin{cases} v_j\ (Good) = 100 \\ v_j\ (Neutral) = 0 \end{cases} \quad (1)$$

and

$$\sum_{j=1}^{n} w_j \ and \ w_j > 0 \ (j = 1,2, \dots, n) \quad (2)$$

and

$$\sum_{i=1}^{m} k_i \ and \ k_i > 0 \ (i = 1,2, \dots, n) \quad (3)$$

Where:

x_j – descriptor of performance of an option on criterion j

$v_j(x)$ – partial score of option x with regard to criterion j

w_j – weighting coefficient of criterion j and k_i – weighting coefficient of area of concern i

n – number of criteria j and m – number of areas of concern i

$v(x)$ – global overall score of option x

The central objective of the additive model is to yield an overall rating of options that depict the actors' preferences elicited throughout the interactive decision-making process. The creation of this model is a recursive process in which feedback is allowed at any stage making the MCDA process versatile and flexible (Bana e Costa *et al*., 1999). In the additive aggregation model, the score appointed to an alternative in a particular criterion may be compensated, either positively or negatively, by the scores acquired in other criteria. The score can also be compensated, affirmatively or opposingly by the scaling constants allocated to each criterion. The additive aggregation, is applied to all areas, permitting the calculation of an overall score of the product. The aggregated scores facilitate in evaluating the global state of the products. The defined category thresholds determine the necessary action to be executed in order to approve or reject the product's aggregated score.

4 Framework Application

The methodological framework proposed in the previous chapter was implemented in the context of developing a sustainability assessment tool. This section aims to present the application procedure and the outcomes of developing the proposed methodology. The results of the Delphi process and application of the constructed model on consumer products are presented later in chapter 5.

4.1 Generation of Potential Criteria

A preliminary task was carried out to generate a list of potential criteria. This stage consisted of 1) an analysis of pre-established PLANETIERS' principles that were used as guidance protocol to evaluate sustainability of products, and 2) a detailed literature review to identify criteria concerning sustainability assessment relevant to household commodities.

The criteria were determined by studying literature relevant to sustainability assessment. The criteria were also based on pre-established principles prior to this dissertation. Instead of coming up with a completely brand-new set of criteria, it was decided to frame some of the criteria based on these principles as they were found to roughly cover the basic characteristics of sustainability with regards to product manufacture and disposal. A certain part of the process was devoted to refining and adjusting the previously devised principles to identify indicators in order to develop the evaluation criteria needed to enable the appraisal of options. The refinement and adjustment procedure began with an analysis of the said 10 principles. The principles proposed by PLANETIERS and their truncated description is presented in Table 1. These principles were based United Nations Sustainable Development Goals, more specifically goal number 12, "*Responsible Consumption and Production*". These principles were generic in nature with no clear distinction and were found to be an agglomeration of many discrete and independent criteria which were needed to be segregated and distinguished. A few of these principles were interlinked and connected and were merged into a single criterion.

Table 1 - PLANETIERS' Principles

Principle	**Description**
Accredited Certification	Possession of valid certificate showcasing social and environmental practices contributing to positive impact
Human Health & Safety	Processes, technologies and materials used in production should not be harmful to human health
Reduced Resource & Energy Consumption	The products manufactured and the processes involved must pertain to reduction in the consumption of resources and energy

Table - PLANETIERS' Principles (Continued)

Principle	**Description**
Circular Economy	Products must have an eco-design that facilitates recycling or reusability thus encouraging sustainable production and avoiding linear consumption
CO_2 Sequestration	The product should contribute to the overall reduction in CO_2 emissions by incorporating CO_2 absorbing materials or encourage renewable energy
Conservation of Natural Resources	The product manufacturer must engage in activities which must mitigate the impact made on the natural environment
Purity in Composition	Raw materials of similar type should be incorporated into the product henceforth avoiding the overuse of diverse set of resources
Preservation of Local Culture & Origin	Products should include local materials and local production processes to safeguard cultural heritage and boost local economy
Social Impact	Products manufactured caters socially excluded people or manufacturing companies should contribute in promoting social welfare
Protection of Environment & Wildlife	Products manufactured or the processes utilize should not be harmful to the local environment and the natural ecosystem

Extensive literature review was carried out on Google Scholar and Science Direct repository, on the subject matter sustainability to identify criteria. Various tools and methodologies for sustainability assessment were reviewed to determine them. The SDGs were studied with respect to acquisition of materials, manufacturing process and disposal of products, which in turn aided in discerning appropriate indicators. Lastly the product portfolio of PLANETIERS was inspected and criteria were ascertained with regards to specification and category. A more pragmatic stance was undertaken in identifying criteria as it was realized, drafting of descriptors of performances would be problematic. Additionally, many Small and Medium Enterprises (SMEs), that constitute majority of the PLANETIERS suppliers, would not have the requisite product data, to allow their assessment

4.2 Delphi Process

Figure 5 shows the process flow during the web-Delphi process. The preliminary stage represents the selection of panel and generating list of potential criteria. It also incorporates the design of the Delphi process, execution of the Delphi survey, and the application of protocol to measure group agreement. The literature study and analysis of the principles, described previously, yielded identification of 34 potential criteria grouped with 3 areas of concern. The list comprised of 23 environmental (67.65%), 5 social (14.70%) and 6 economic (17.65%) criteria. The panel was requested to review these criteria and indicate their level of agreement or disagreement with regards to their relevance in appraising sustainability of PLANETIERS' product portfolio. The most commonly used type of response scale in Delphi study surveys is a rank-ordered or Likert-type scale (Belton *et al*., 2019), and this scale was employed in the eliciting the panel's response. Each potential criterion was supplemented with a short description pertaining its calculation, data availability and accessibility. The structure and format of questionnaires directly influence the nature of the responses obtained from panelists (Belton *et al*., 2019), thus the criteria were solicited in a sequential manner. Panelists were advised to share their views and comments about the potential relevance of an criteria and any perceived difficulties associated with it in collecting or processing the required data.

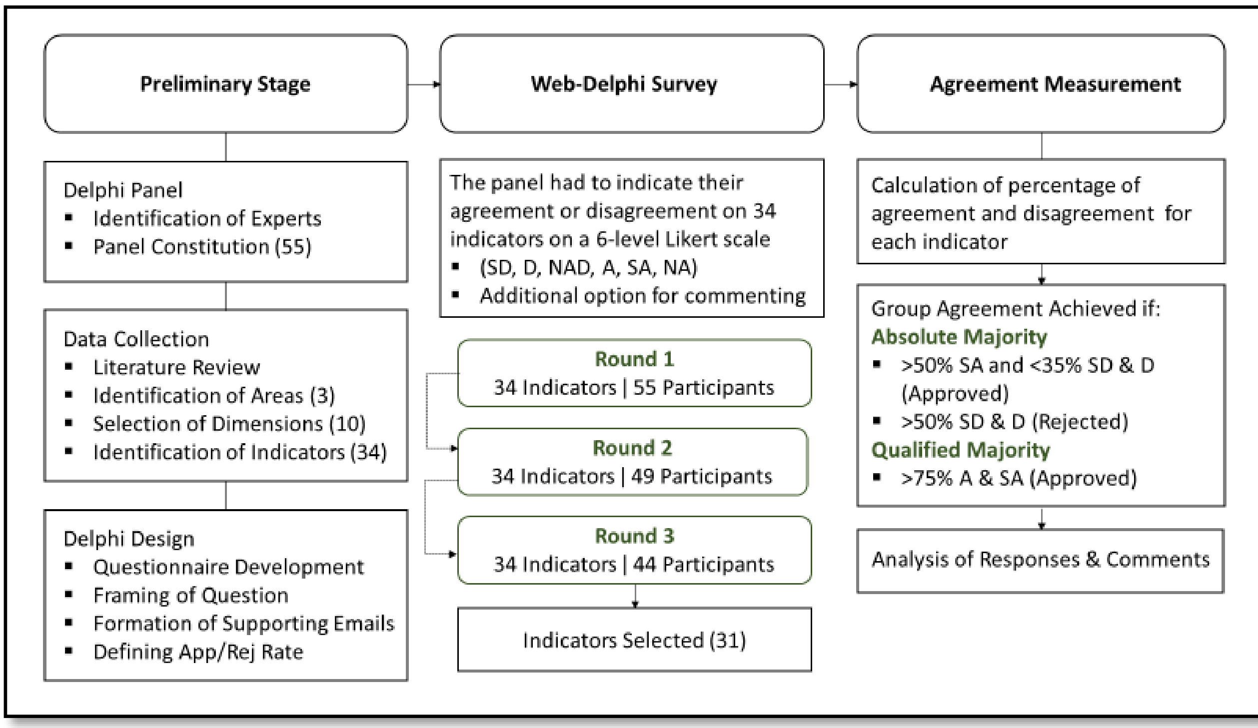


Figure 5 - Web-Delphi Process Flow Diagram

In this preparation stage for the Delphi process, supporting letters were prepared and adjusted according to the number of iterative questionnaire sessions. Before sending a formal invitation, experts were inquired about their willingness to participate in the Delphi process as part of the panel. This informal inquest culminated in a high participation rate across the three rounds.

4.2.1 Delphi Panel

There were in total 55 participants including 6 stakeholders and 49 experts. The Delphi panel reflected a wide range of expertise profiles: Stakeholders (6 panelists); Energy Management (9 panelists); Environmental Engineers and Sustainability Consultants (21 panelists); Professors (4 panelists); Climate Activists, Journalists and Politicians (6 panelists); Waste Management (1 panelist); Entrepreneurs and Sellers (4 panelists) and Engineers and Scientists (4 panelists). The panel reflects diversity with respect to different geographical representation: Southern Europe (37 from Portugal), Western and Eastern Europe (4 from Macedonia, Poland and Germany), Latin America (2 from Brazil) and Middle East and South East Asia (12 from Syria, Saudi Arabia and Pakistan).

4.2.2 Delphi Application

The Delphi survey was employed by making use of the "ENERPHI" project, a platform for participatory processes in decision-making contexts established exclusively for students (ENERPHI, 2019). The Delphi process was conducted via a web platform "WELPHI", (https://app.welphi.com/welphi/Pages/LoginPage.aspx), for implementing and monitoring the participatory process involved in this dissertation. WELPHI allows facilitators to implement the Delphi method in their own projects whilst retaining the Delphi's main features (ENERPHI, 2019). It is reported that three rounds are optimal and consisting fewer than three rounds would not induce stability in results. Henceforth, the survey was delivered across the three rounds of study, parallel to questionnaire reminders and providing feedback through the web platform.

The experts who had formerly stated their willingness to participate were dispatched an online invitation. This invitation email contained individual credentials to log in to the web platform alongside information regarding the study design and detailed instructions concerning the first round. This is their first interaction with the Welphi platform. Reminders were sent to those participants who did not complete the survey in the previously specified time-frame. (Hasson et al., 2000) mentions "the type of communication techniques employed can have a profound impact on the results obtained." Giving panelists repeated reminders by email and text message helps to achieve a high response rate (> 90%) throughout the Delphi process (Belton *et al*., 2019), and for that reason reminders were sent through the web platform. Additionally, direct reminders were sent through informal channels which had been previously used for contacting the experts. The application of multiple rounds, allows forth stability and consistency of the answers from the respondents. This accounted for a very high participation rate. The panel members who did not respond to a round, within the time period and after deadline extension, were not invited to partake in further rounds.

- In this first round, panelists were required to indicate their agreement or disagreement with the following statement "*Is this indicator relevant in assessing the sustainability of a product?*" Description for each one of the potential criteria was available under the “Eye” icon throughout the three rounds. The panel would response on a 6-level Likert scale, with Disagree (D) and Strongly Disagree (SD) indicating disagreement and Agree (A) and Strongly Agree (SA) indicating agreement. Moreover, panelists could indicate their uncertainness by opting Neither Agree nor Disagree (NAD). Furthermore, if the panelist did not wish to answer or believed that certain criterion did not fall under their expertise, they could select the option No Answer (NA). This was in accordance to (Trevelyan and Robinson, 2015) suggestion of four to seven being the optimal number of categories on the Likert scale. Additionally, dialog box was provided to participants where they could voice their comments. Once the participants had given their answers before the prescribed deadline, the 1st round was terminated.

- An invitation to take part in this 2nd round was sent by e-mail. For numerical measures such as Likert-type scales, the reporting of central tendencies at the feedback stages of a Delphi study is beneficial as it allows panelists to see how their response compares to the group response as a whole (Hasson *et al*., 2000). In the second round, panelists were presented with the results of the first round, an anonymous aggregated sum of votes for each criterion on each Likert scale option. They could also view their own response from the previous round represented in the shaded region. The aggregation of all individual answers for each criterion is showcased next to the radio button for each Likert item. Panelists had the possibility of viewing comments from the previous round. Participants still engaging in the process had the option to revise their previously given answers. This enabled each panelist to consider his or her own answer with respect to the group opinion and could either maintain their answer or change their answer from the previous round by clicking the “EDIT” icon. The 2nd round of the Delphi process was ended once the deadline had passed.

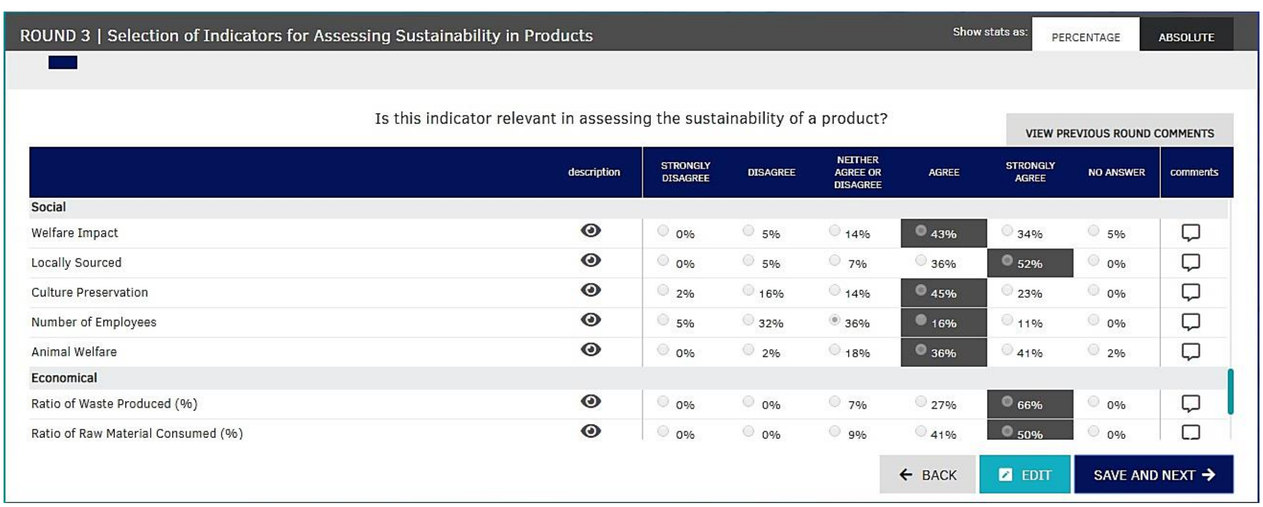


Figure 6 - Screenshot of 3rd Round

- Once again, invitation to the 3rd round of the questionnaire was sent via e-mail. In the third round, the same procedure was replicated as in the second round. Similarly, participants were invited to revise their previously given answers. The screenshot of the 3rd round is displayed in Figure 6. Again, comments could be provided wherever participants saw fit. This ended the 3rd and final round of the survey and henceforth concluded the Delphi process.

After the three rounds had concluded, a final report with the results of the Delphi process was sent to participants. This report contained the final aggregated answers provided from all the participants from the 3rd round, highlighting the criteria where consensus was achieved either in agreement or disagreement. The supporting emails for the complete Delphi process are listed in Appendix A.

The Delphi process was carried out between July and August 2019. The interval between rounds was deliberately kept small to facilitate consensus. As greater the timespan between rounds, the more changes in individuals' viewpoints and more divergent the group responses. Its detailed schedule is shown in Table 2 including the starting, duration and ending dates of the three rounds of the web-Delphi process. Panelists were asked to answer within 1 week since the commencement of the survey which was instigated on dispatching the e-mail. Deadline to the questionnaire was extended, for the first round, to give participants more time to engage in the process. In the subsequent rounds, sizeable response rate was achieved to render deadline extension unnecessary.

Table 2 - Delphi Schedule

	Start	**End**	**Duration**	**Extension**
1st Round	18th July 2019	24th July 2019	7 days	1 day
2nd Round	26th July 2019	31st July 2019	6 days	0 days
3rd Round	2nd August 2019	7th August 2019	6 days	0 days

The group consensus was defined by calculating the percentage of responses, given in each Likert item, for each criterion in each round. Group consensus was obtained at the end of the third round by calculating the majority of agreements and disagreements expressed by the participants' answers "SA", "A" and "SD" and "D" in percentages per criterion. Criteria receiving more than 50% of "SA" responses and at the same time, did not receive more than 33.3% of "SD" and "D" responses, were approved by *absolute majority*. Criteria with more than 50% of either "SD" or "D" responses were rejected by *absolute majority*. If an criterion was neither approved or rejected by *absolute majority*, then upon receiving more than 75% of "SA" and "A" responses was considered approved by *qualified majority*. This agreement protocol was decided by the PLANETIERS team and is depicted in Figure 7.

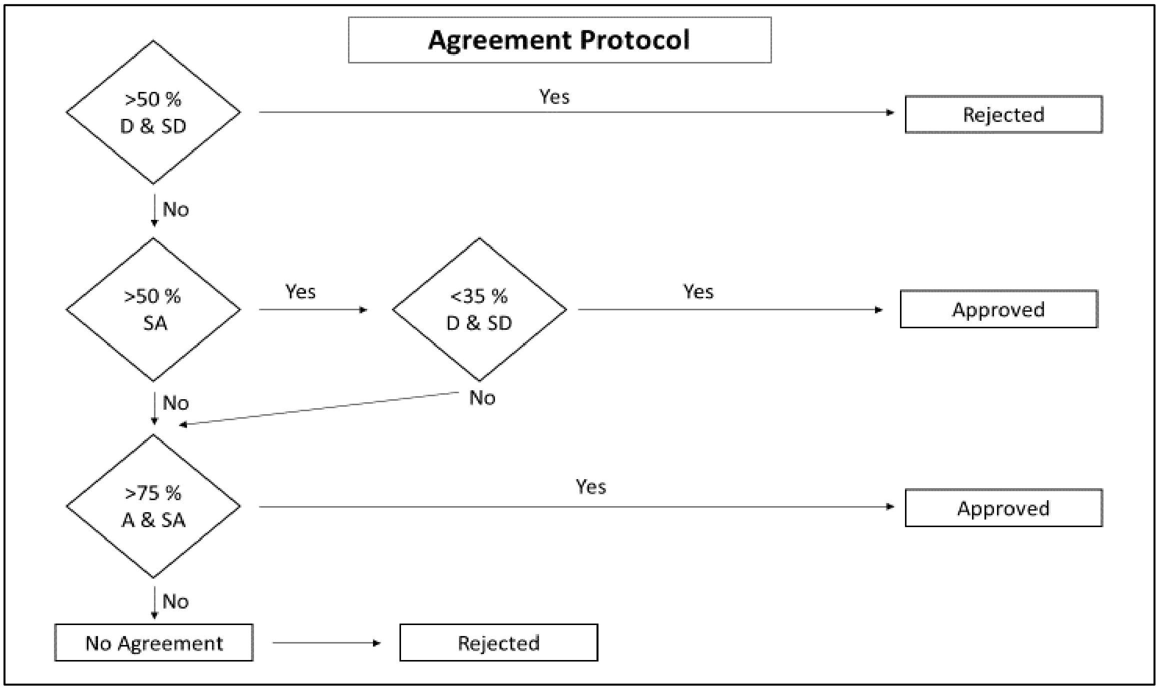


Figure 7 - Agreement Protocol

The answers of the questionnaire were discussed with the stakeholders and were used to identify and finalize the criteria for the construction of the multicriteria model. The compiled responses of the participants enabled the construction of the multicriteria evaluation model with the aid of the decision support software M-MACBETH in determining the criteria value functions and their weights for the generation of the additive aggregation model.

4.3 Model Implementation

At this stage the results of the Delphi process were obtained and the next step was to build the multicriteria evaluation model. It was necessary for the development of the product assessment tool. A three-day decision conference was exercised and the PLANETIERS team was involved in the interactive process. The decision makers during the structuring and evaluation stages of the model building process were Pedro Carreira, Madalena Ferreira, Diogo Lourenco and Teresa Kuski as part of the PLANETIERS team. The author of the thesis acted as an impartial facilitator. An overhead projector screen and a large screen monitor were utilized in the process. The M-MACBETH decision support system, operated using a computer, was utilized for the application of the model and connected to the projector. The large monitor, connected to a second laptop, was used to show the list of evaluation criteria together with their definitions and reference levels. M-MACBETH enables the structuring of value trees, construction of criteria descriptors, scoring of performances and formulation of a value scale, weighting of criteria, and perform sensitivity and robustness analysis of the inherent value of options (Bana e Costa *et al*., 2003).

4.3.1 Screening Criterion

The purpose of the screening criterion is to prequalify or shortlist the alternatives for evaluation. They can also be implemented to deliberately comply with the specific requirements of the decision maker or the organization. In the case of PLANETIERS, the screening criterion was based on the principle "Human Health and Safety". A discrete, concise and clear descriptor was assigned and made operational as the screening criterion. This criterion became the threshold for admissibility for the products to proceed towards appraisal. It is important to notify that as of yet, no such product has been incorporated by PLANETIERS which may be a subject to human safety. The screening criterion is mentioned below:

❖ "Toxicity in Product Manufacture and Waste" – The manufacturing of the product does not contain any toxic materials and the waste generated is not toxic.

4.3.2 Value Trees

The results of the Delphi process were needed to be analyzed for the creation of the Value Tree. On analysis of the identified criteria, it became apparent that the criteria would be defined in a hierarchical structure, based on the three pillars of sustainability. It was further deduced that the Environmental pillar would need to be further split to represent each stage of the product manufacturing process. The criteria were also scrutinized on the basis of mutually preferential independence. This resulted in reduction in the number of approved indicators utilized as criteria for the model. An additional criterion was added during the decision conference when a particular value concern was not captured in the initial set of criteria. The criterion "Biodegradable" was included so that all the essential value concerns of decision-makers could be addressed. Once the analysis was complete, it was possible to represent this information in the form of a value tree in which criteria are grouped according to their area of concern. A holistic view of the model structure is obtained through the creation of the value tree.

The Value Tree, a hierarchical structure, takes the form of a generic value tree embodying the evaluation criteria for appraising products. The Value Tree, depicted in Figure 8, consists of three areas of concern, comprising the three domains of sustainability. The generic value tree was produced with three criteria domains and a total of 23 criteria, each operational by respective descriptors of performance. The value tree aims to capture a comprehensive generic set of value concerns that can be adapted to different product categories.

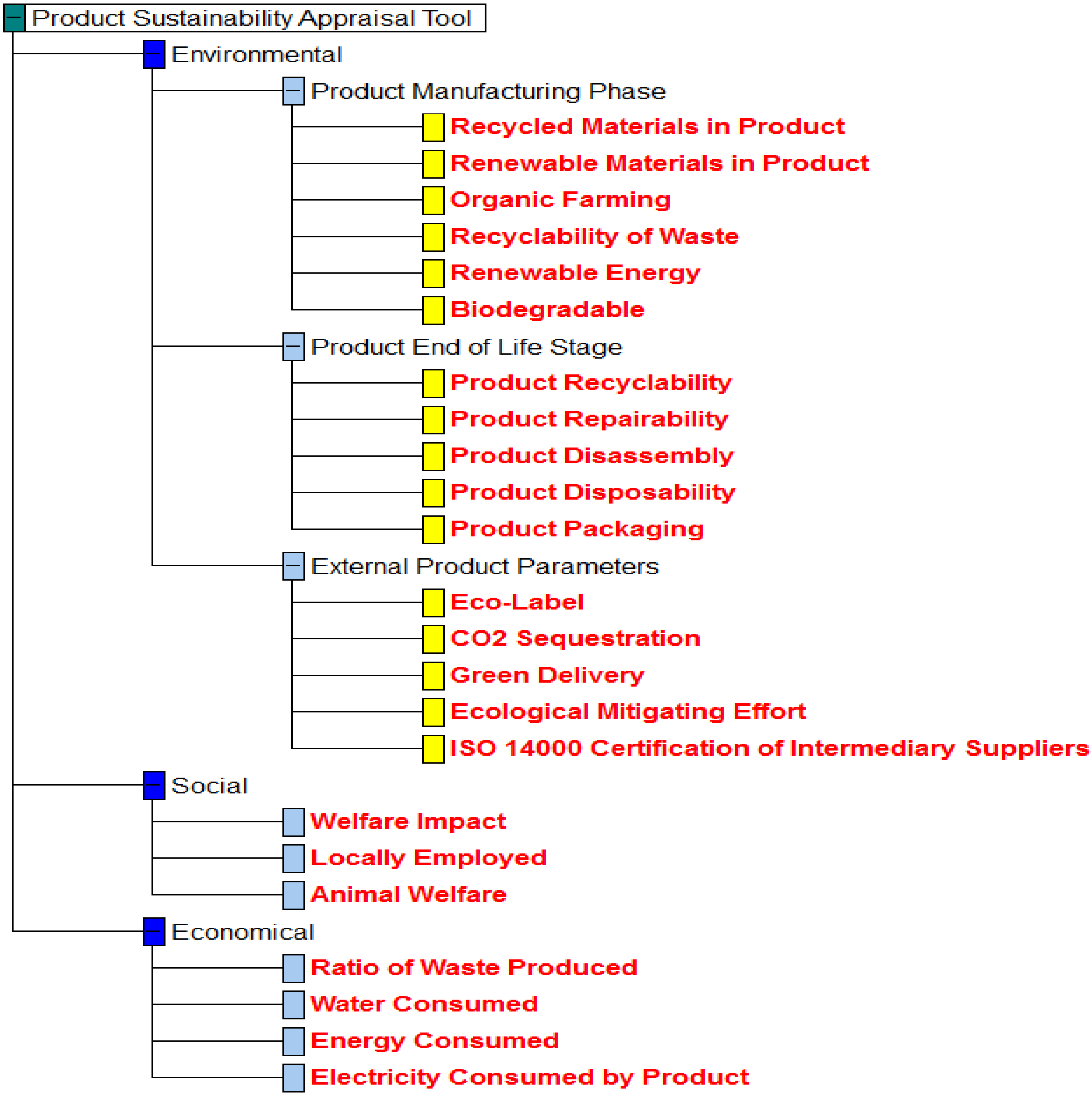


Figure 8 - Generic Value Tree - "Product Sustainability Assessment Tool"

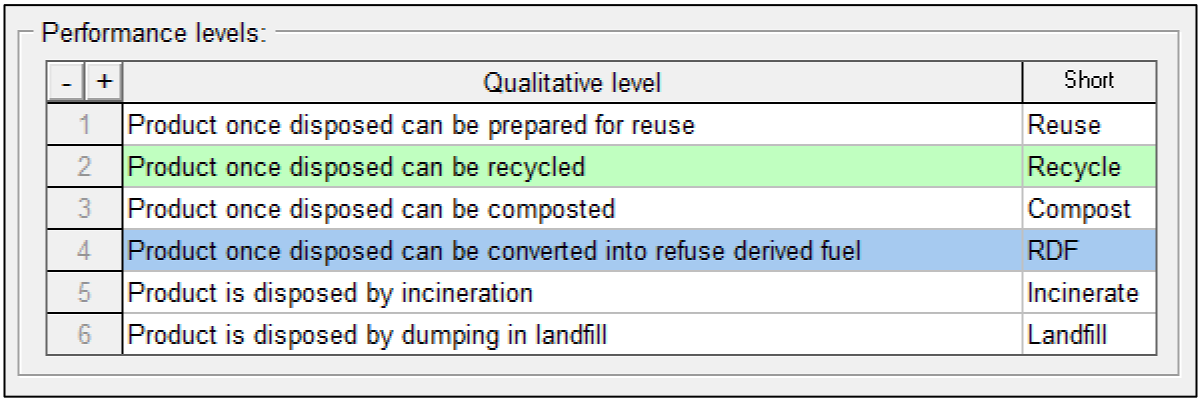
The generic form of Value Tree was adapted into a product category specific value model using a bottom-up approach by analyzing the characteristics of the product categories to be evaluated. The PLANETIERS team were consulted in this adaptation process as they possessed in-depth knowledge of the product portfolio. This was necessary as several criteria identified were relevant and pertinent to one particular category but inappropriate to the other. These criteria would have affected the result of the model hence were excluded from some models. The participants considered these criteria being non-fundamental for the evaluation of specific category of products and would be considered as equivalent of being assigned a zero weight. The results of the generic model would have created disagreements on this basis and “lack of trust” with the results. The evaluation criteria in the generic model were utilized as criteria for the models of specific categories, as they were able to capture the performance levels of the alternatives and possessed the required properties such as non-redundancy and preferential-independence to ensure the creation of an adequate evaluation model. The weights of each category model were assigned through the same approach as mentioned in section 3 (sub-section 3.5.2). It was decided that the product evaluation models would be created for 4 specific categories; Fashion, Hygiene, Home Décor and Reusables. The product category specific value trees were validated by the decision conference participants through the model building process and after testing the model on the products.

The product category specific Value Trees are shown in Appendix B.

4.3.3 Descriptors of Performance

A criterion is operational from its prospect to be assigned a descriptor of performance. Variety of evaluation criteria necessitate the requirement of different combination types of descriptors. These contrasting types of descriptors of performance need to be taken under consideration that best suit the evaluation criteria. The decision maker defines plausible levels of performance to accommodate impacts of different alternatives. Determining descriptors of performance may also reveal presence of redundant criteria, in such case, the model should be revised, which may require merging or elimination of certain criteria. If required, descriptors of performance of multiple criteria may be clustered into a single evaluation criterion. This case was applicable for a few criteria amongst them was “Organic Farming” which involved merger of the two criteria into one criterion. This recursive development process simplified the criterion and validated the soundness of the model.

A six-level descriptor of performance was built for the criterion ‘‘Product Disposability’’. This descriptor was constructed in three steps: (1) The stakeholders were invited to describe the “Best” and “Worst” levels of performance (“Reuse” and “Landfill” respectively). (2) Four intermediate levels (Recycle, Compost, RDF and Incinerate) were defined. (3) The stakeholders further had to identify which of the levels corresponds to the “Good’’ and “Neutral” reference levels. The “Good” level (Recycle) was always associated with the reference level that was considered desirable and the “Neutral” level (RDF) was aligned with the performance level that was considered neither bad nor favorable. An illustration of the qualitative descriptor of performance regarding “Product Disposability” is shown in Figure 9.



Performance levels:

- +	Qualitative level	Short
1	Product once disposed can be prepared for reuse	Reuse
2	Product once disposed can be recycled	Recycle
3	Product once disposed can be composted	Compost
4	Product once disposed can be converted into refuse derived fuel	RDF
5	Product is disposed by incineration	Incinerate
6	Product is disposed by dumping in landfill	Landfill

Figure 9 - Descriptors of Performance for the criterion "Product Disposability"

Table 3 defines the list of all criteria and their descriptions.

Table 3 - List of Criteria and their descriptions

Criteria	**Description**
Recycled Materials in Product	Raw material used in the manufacture of product are obtained through recycling
Renewable Materials in Product	Materials used in the manufacture of the product which can be remade, regenerated or regrown
Organic Farming	The use of ingredients and raw materials in the manufacture of product harvested through organic farming practices
Recyclability of Waste	Waste generated in the manufacture of the product can be recycled
Renewable Energy	Renewable energy consumed in the manufacture of the product
Biodegradable	The use of biodegradable raw materials in the manufacture of product
Product Recyclability	Degree of recyclability of product after it has reached its end of life
Product Repairability	The level of technical expertise required to repair the product
Product Disassembly	The level of technical expertise required to disassemble the product
Product Disposability	The disposal method of the product within waste management hierarchy
Product Packaging	The utilization and disposal of product packaging
Eco-Label	Accreditation of the Product
ISO 14000 Certification of Intermediary Suppliers	ISO 14000 Certification of Suppliers providing raw materials to the product manufacture
CO_2 Sequestration	Utilization of raw materials that have captured CO_2
Green Delivery	The use and disposal of packaging utilized in logistics of the product

Table 3 - List of Criteria and their descriptions (Continued)

Criteria	**Description**
Ecological Mitigation Effort	The manufacturing company of the product partakes in activities to minimize or reverse its impact on the environment
Welfare Impact	The product manufacturing company partakes in activities that promotes social and welfare development
Locally Employed	The manufacture of product promotes local employment and use of local resources
Animal Welfare	The manufacture of product involves testing on animals
Ratio of Waste Produced	The ratio of waste produced to the product manufactured.
Water Consumed	Impact of product manufacture on water consumption
Energy Consumed	Impact of product usage on energy consumption
Electricity Consumed by Product	Type of energy source utilized in the consumption of electricity by the product

The descriptors of all criteria are presented below from Table 4 through 26. All levels of performances are showcased in deceasing order of attractiveness.

Table 4 - Constructed Performance Scale for "Recycled Materials in Product" (Environmental)

Recycled Materials in Product	
Very Good	The product is manufactured from completely recycled materials
Good	More than half of the product is manufactured from recycled materials
Better	Less than half of the product is manufactured from recycled materials
Neutral	The product is not manufactured from recyclable materials

Table 5 - Constructed Performance Scale for "Renewable Materials in Product" (Environmental)

Renewable Materials in Product	
Very Good	The product is manufactured completely from renewable materials
Good	The product is manufactured partially from renewable materials
Neutral	The product is not manufactured from renewable materials

Table 6 - Constructed Performance Scale for "Organic Farming" (Environmental)

Organic Farming	
Good	The product manufactured has all ingredients obtained from organic farming practices
Better	The product manufactured has at least one but not all ingredients obtained from organic farming practices
Neutral	The product manufactured has no ingredients obtained from organic farming practices

Table 7 - Constructed Performance Scale for "Recyclability of Waste" (Environmental)

Recyclability of Waste	
Good	The waste generated in the manufacture of the product is completely recycled
Better	The waste generated in the manufacture of the product is partially recycled
Neutral	The waste generated in the manufacture of the product cannot be recycled

Table 8 - Constructed Performance Scale for "Renewable Energy" (Environmental)

Renewable Energy	
Very Good	The product was manufactured only using renewable energy
Good	The product was manufactured with both renewable and conventional energy
Neutral	The product was not manufactured with renewable energy

Table 9 - Constructed Performance Scale for "Biodegradable" (Environmental)

Biodegradable	
Very Good	Product is made completely from bio-degradable materials
Good	Product is made partially from bio-degradable materials
Neutral	Product is not made from bio-degradable materials

Table 10 - Constructed Performance Scale for "Product Recyclability " (Environmental)

Product Recyclability	
Very Good	The product is completely recyclable
Good	More than half of the product is recyclable
Better	Less than half of the product is recyclable
Neutral	The product is not recyclable

Table 11 - Constructed Performance Scale for "Product Repairability " (Environmental)

Product Repairability	
Very Good	The product can be repaired by the consumer without product manual
Good	The product can be repaired by the consumer through the use of product manual
Better	The product can be repaired by community hardware stores
Moderate	The product can only be repaired by supplier
Neutral	The product does not need to be repaired
Bad	The product cannot be repaired

Table 12 - Constructed Performance Scale for "Eco-Label " (Environmental)

Eco-Label	
Good	Product manufacturing company possesses environmental performance certification
Neutral	Product manufacturing company does not possess environmental performance certification

Table 13 - Constructed Performance Scale for "Product Disassembly " (Environmental)

Product Disassembly	
Very Good	The product can be disassembled at the end of life by the consumer without use of tools
Good	The product can be disassembled at the end of life by the consumer with use of common tools
Better	The product can be disassembled at the end of life by third party with use of specific tools
Neutral	The product does not need to be disassembled
Bad	The product cannot be disassembled at the end of life

Table 14 - Constructed Performance Scale for "Product Disposability " (Environmental)

Product Disposability	
Very Good	Product once disposed can be prepared for reuse
Good	Product once disposed can be recycled
Better	Product once disposed can be composted
Neutral	Product once disposed can be converted into refuse derived fuel
Bad	Product is disposed by incineration
Worse	Product is disposed by dumping in landfill

Table 15 - Constructed Performance Scale for "Product Packaging" (Environmental)

Product Packaging	
Very Good	Product does not have any packaging
Good	Packaging of the product can be reused multiple times and can be recycled at the end of use
Better	Packaging of the product can be reused multiple times but cannot be recycled at the end of use
Moderate	Packaging of the product can be recycled and is biodegradable
Neutral	Packaging of the product can be recycled but is not biodegradable
Bad	Packaging of the product cannot be reused or recycled

Table 16 - Constructed Performance Scale for "Green Delivery" (Environmental)

Green Delivery	
Very Good	Delivery Packaging is made out of recycled materials and can be reused
Good	Delivery Packaging can be reused but is not made out of recycled materials
Better	Delivery Packaging cannot be reused but is made out of recycled materials
Neutral	Delivery Packaging cannot be reused and is not made out of recycled materials

Table 17 - Constructed Performance Scale for "CO2 Sequestration" (Environmental)

CO2 Sequestration	
Very Good	Products manufactured in final form or intermediate form utilize Cork as raw material which is harvested in a sustainable manner
Good	Products manufactured in final form or intermediate form utilize Cork & Wood as raw materials which are harvested in a sustainable manner
Better	Products manufactured in final form or intermediate form utilize Wood as raw material which is harvested in a sustainable manner
Moderate	Products manufactured in final form or intermediate form utilize plant-based materials besides Cork & Wood as raw materials which is harvested in a sustainable manner
Neutral	Products manufactured in final form or intermediate form do not utilize Cork, Wood or other plant-based materials as raw materials

Table 18 - Constructed Performance Scale for "Ecological Mitigation Effort" (Environmental)

Ecological Mitigation Effort	
Good	Product manufacturing company directly engages in activities to mitigate ecological impact
Better	Product manufacturing company engages in activities through an external party to mitigate ecological impact
Neutral	Product manufacturing company does not engage in activities directly or through an external party to mitigate ecological impact

Table 19 - Constructed Performance Scale for "ISO 14000 Certification of Intermediary Suppliers " (Environmental)

ISO 14000 Certification of Intermediary Suppliers	
Good	Intermediary suppliers of the product are ISO 14000 series certified
Neutral	Intermediary suppliers of the product are not ISO 14000 series certified

Table 20 - Constructed Performance Scale for "Welfare Impact" (Social)

Welfare Impact	
Good	Product manufacturing company engages in activities towards social development or welfare support
Neutral	Product manufacturing company does not engage in activities towards social development or welfare support

Table 21 - Constructed Performance Scale for "Locally Employed" (Social)

Locally Employed	
Good	Product is manufactured by employing local people and resources exclusively within a single district zone
Better	Product is manufactured by employing people and resources within a national region
Neutral	Product is manufactured by employing people and resources from both domestic and international regions
Bad	Product is manufactured by employing people and resources exclusively from international regions

Table 22 - Constructed Performance Scale for "Animal Welfare" (Social)

Animal Welfare	
Good	Product manufacturing process does not involve mal treatment of animals and does not involve testing on animals
Neutral	Product manufacturing process does not involve interaction with animals
Bad	Product manufacturing process involves testing on animals but does not involve mal treatment of animals

Table 23 - Quantitative Performance Scale for "Ratio of Waste Produced" (Economical)

Ratio of Waste Produced	
Very good	1
Good	10
Better	20
Neutral	30
Bad	40

Table 24 - Constructed Performance Scale for "Water Consumed" (Economical)

Water Consumed	
Good	The manufacture of product helps to reduce water consumption
Neutral	The manufacture of product does not impact the consumption of water

Table 25 - Constructed Performance Scale for "Energy Consumed" (Economical)

Energy Consumed	
Good	The use of product helps reduce energy consumption
Neutral	The use of product does not impact the consumption of energy

Table 26 - Constructed Performance Scale for "Electricity Consumed by Product" (Economical)

Electricity Consumed by Product	
Good	Electricity consumed during charging or operation is obtained from renewable methods
Neutral	No electricity is required either for charging or operational use
Bad	Electricity consumed during charging or operation is obtained from conventional methods

4.3.4 Value Functions

For the construction of value functions, the reference performance levels needed to be assigned. This step was completed in the previous section by assigning "Good" and "Neutral" levels. Most of the "Neutral" references corresponded with least ranked performances levels due to the ambiguity of sustainability and to the generic nature of the model. It is important to note, similarly to the descriptors, the same value scales are being used throughout the 4 evaluation models to create homogeneity amongst the models. The "Green" and "Blue" cells represent "Good" and "Neutral" reference levels respectively.

Product Disposability

	Reuse	Recycle	Compost	RDF	Incinerate	Landfill	Current scale
Reuse	no	strong	strong	v. strong	v. strong	extreme	150
Recycle		no	strong	strong	v. strong	v. strong	100
Compost			no	strong	strong	strong	50
RDF				no	weak	strong	0
Incinerate					no	moderate	-40
Landfill						no	-89

Consistent judgements

Figure 10 - Judgement Matrix for the "Product Disposability"

The judgement matrix is filled by first requesting the stakeholders to rank the performance levels in terms of attractiveness from the "most" to "least" attractive. After the performance levels have been ranked, the decision makers were asked to express the magnitude of difference of attractiveness between consecutive performance levels. The question was phrased as, "Given two performance levels, with the first better than the second, describe the difference of attractiveness (value) between them is?" The panel had the option to respond from the seven semantic judgements; "*No*", "*Very Weak*", "*Weak*", "*Moderate*", "*Strong*", "*Very Strong*" and "*Extreme*". The assessor is inquired to choose one of the categories to facilitate pairwise comparison. Figure 10 depicts the judgement matrix for "Product Disposability" criterion.

For quantitative descriptors, a graphical piecewise linear function is generated. This is showcased in figure 11 for the criterion "Ratio of Waste Produced".

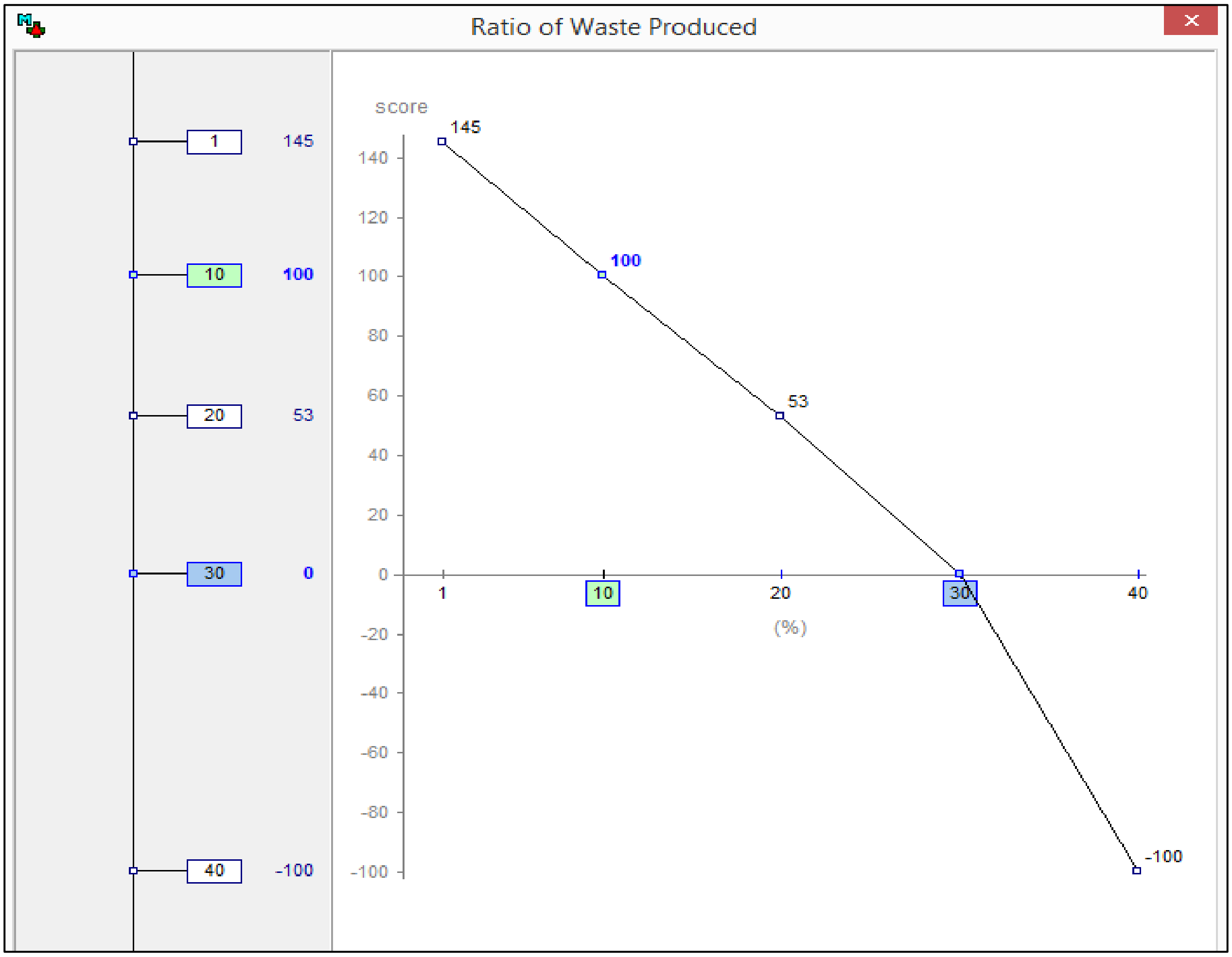


Figure 11 - Graphical representation of the value function for the criterion "Ratio of Waste Produced"

A value scale was attributed with each performance scale of each criterion, enabling a value score to be assigned to the product's performance, which capacitated the measurement of the product's attractiveness at each criterion. The generated value functions were validated by the decision makers. Value functions of the all the criteria are in Appendix C. After all criteria have been assigned a value sale, the next step in the decision conference was to assign weights to the evaluation criteria.

4.3.5 Weighting Criteria

The weighting of criteria was solely based on the "Good" and "Neutral" references levels defined for the criteria. The hierarchical structure of the value tree required the usage of the hierarchical weighting scheme. A bottom-up approach was utilized, where first the criteria in a single area were weighted, then the areas were weighted with respect to hierarchical structure.

- The first step pertains to weighting of the criteria within the same area. The stakeholders were requested to rank the swing on each individual criterion in terms of decreasing order of attractiveness. They were told to assume that each criterion was at “Neutral” level and asked to select which criterion would they prefer most to improve from “Neutral” to “Good” performance level. This would be repeated until the ranking of swings on all criteria is complete. The appraiser is asked the importance of improving the performance of that criterion with reference to the last column of the judgement matrix labelled “Neutral”. It is inquired *“For each criterion, the importance of improving the performance from the “Neutral” to the “Good” level is?”* The seven semantic categories are available to choose from. They will then be inquired to compare consecutive criteria and share the degree of importance of improving from “Neutral” to “Good” of performance level on each consecutive criterion. The following question was phrased, *“For each criterion, both acquiring “Neutral” performance level were to improve to “Good” performance level, describe the difference of attractiveness (value) between them is?”* Again, the panel were given the option to select from the seven semantic judgements. Once the judgement matrix is complete, an interval scale is generated. The “Neutral” level and the most preferred criterion are fixed at “0” and “100” respectively, the criteria between them were adjusted till the decision makers were satisfied, whilst maintaining the compatibility with the judgement matrix. The M-MACBETH software generates the weight of the criteria through normalization of the scores on the interval scale.

- The next step is to weigh the areas in order to harmonize the weights of the criteria. Under the “Environmental” aspect of sustainability, the criteria are allocated to three areas with regards to “Manufacturing Phase”, “Product End of Life” and “External Parameters”. The panel was asked, *“Consider the criterion in each area with the highest weightage, with both criteria being in their “Neutral” reference levels, which one would you select first to move to the “Good” reference level?”* The panel indicated their most preferred criterion. They are then asked to express their desirability in improving the criterion from “Neutral” to “Good” performance level in comparison to the other. The panel was inquired, *“How much more desirable is the improvement in first criterion, such that the first is more attractive than the second?”* The panel were given the option to select from the seven semantic judgements.

- The final step is to weigh the three areas of concern. Exactly the same process is repeated and the same weighting procedure is applied. The question is phrased again and the decision makers then respond with their selected criteria amongst the *Environmental*, *Social* and *Economical* areas. After all the areas are weighted, the software allows to view the weights of each criteria simultaneously. These weights are the global representation of the criterion on the additive model. The actors were asked to adjust the value of weights as long it does not result in the change in the ranking of the criteria. This concludes the weighting procedure and facilitates an area value score to be assigned to a product.

Figure 12 shows the overall weightages of all criteria for the generic model. The product category specific weights of all criteria for the 4 evaluation models is presented in Table 27.

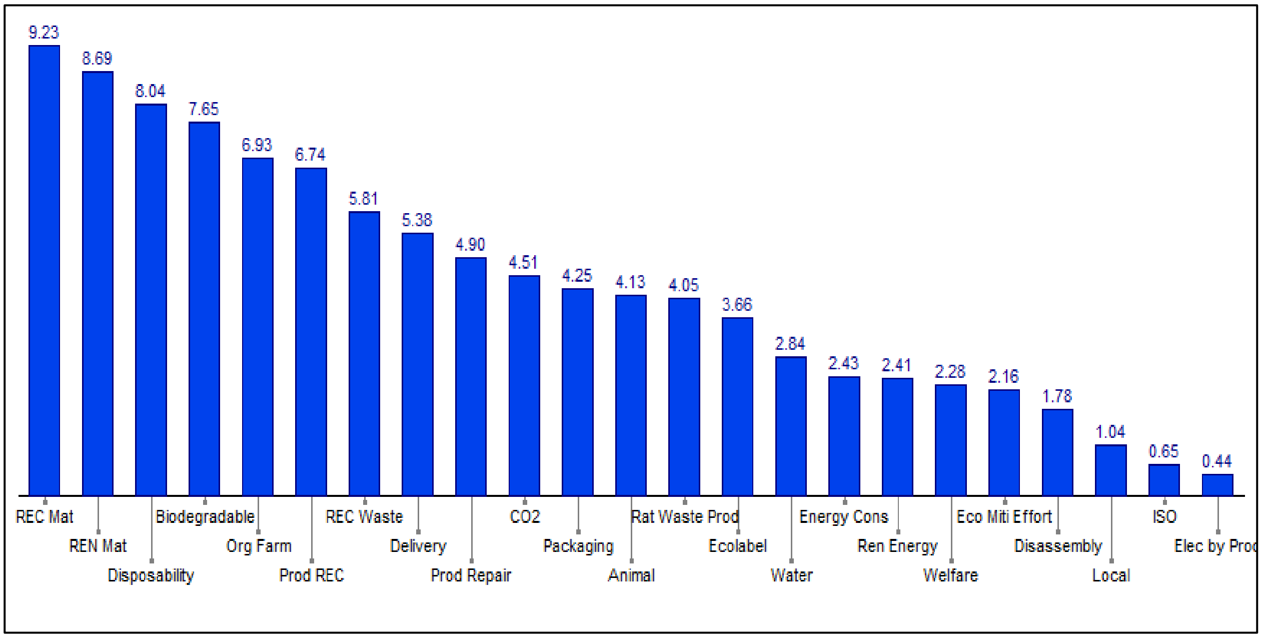


Figure 12 - Generic Model Weights "Product Sustainability Assessment Tool"

The assigning of weights completes the construction of the additive model and ends value model building stage. The validity of the models generated is attained from performing the sensitivity and robustness analysis. Once the model structure has been defined, criteria identified, descriptors constructed, value functions generated, weights assigned and the additive model established, the alternatives can be evaluated.

4.3.6 Category Thresholds

To analyze the results, the category thresholds (depicted in Figure 2), were established for the classification of products upon assessing sustainability of products. As there being an infinite number of products, the consequent results would be covering over a broad range of scores. The top and bottom thresholds were based on the “Good” and “Neutral” levels. Alternatives scoring above the level “Good” (possessing a score greater than 100) would be automatically approved as showcased by fictitious product “**A**” on category threshold scheme (Figure 14). Alternatives scoring below the level “Neutral” (possessing a score lower than 0) would be automatically rejected as showcased by fictitious product “**D**” on category threshold scheme. An issue arose in the intermediary range between “0” and “100” as this range would occupy a large number of products and how to handle the products within this range. This range was split into two categories, “Monitoring Zone” and “Grey Zone”. However, the threshold between these two zones had to be established. The approach utilized in (Bana e Costa and Oliveira, 2002), was implemented to determine the threshold based on the identification of “*reference profiles*”. This is defined in terms impact levels on all criteria for which the decision makers hesitated on which category the product should be assigned. The identification of “*reference profile*” is conducted by *Bottom-Up procedure* as represented in Figure 13. The *Bottom-Up* procedure consists of the following stages:

1. A fictitious alternative is characterized by Neutral level in all criteria.
2. At each stage, a criterion is selected and the level assigned is raised from “Neutral” to “Good”
3. This is continued until decision maker assigned the product from Grey Zone to Monitoring Zone
4. The, the level of the last criterion was lowered, until decision makers hesitated in assigning a category

This was computed by the evaluation model using the M-MACBETH software and resulted in the score of “50”. This score defined the threshold for separating the two categories.

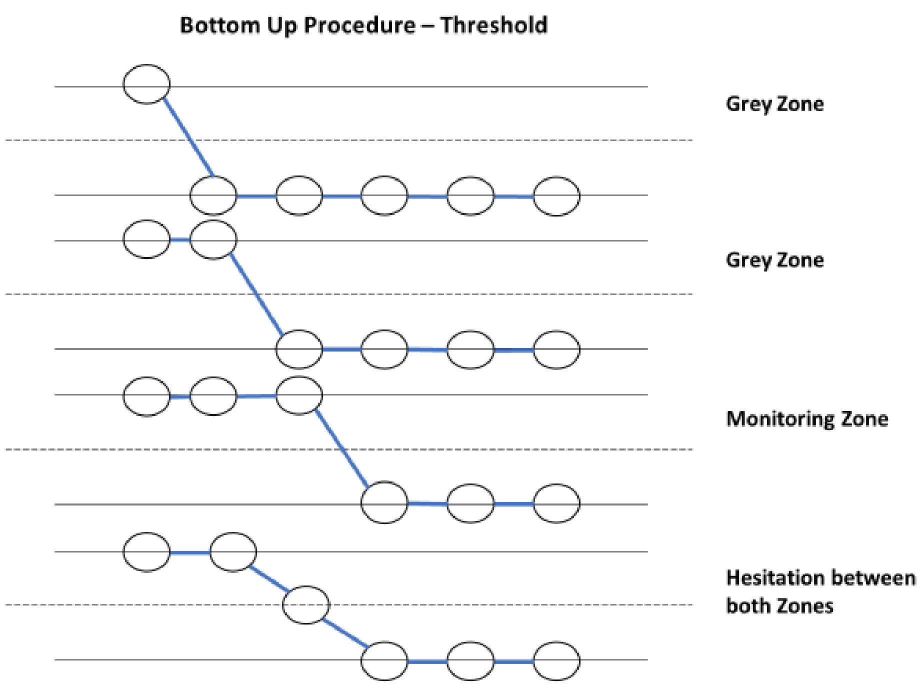


Figure 13 - Bottom-Up procedure to identify reference profile

The products scoring between “50” and “Good” would be approved for the online platform but will be monitored by the team as showcased by fictitious product “**B**” on category threshold scheme. The products scoring between “Neutral” and “50” would be scrutinized and assessed over certain criteria depending on the specific evaluation model. If a performance level of the product in the particular criterion is either at “Good” or a higher level, then the product is approved (otherwise rejected) as showcased by fictitious product “**C**” on category threshold scheme. The category threshold scheme is represented below in Figure 14.

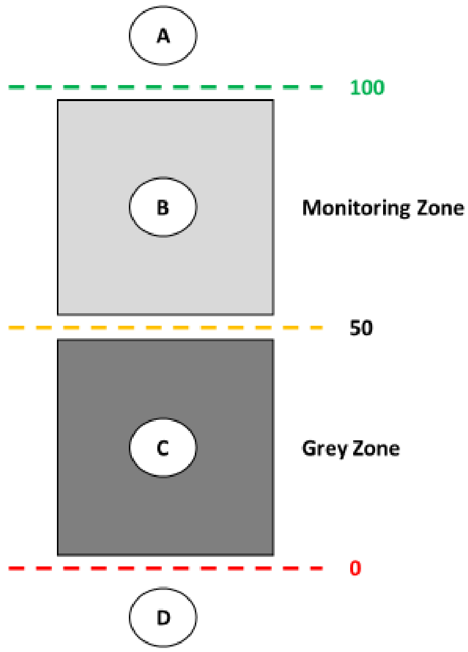


Figure 14 - Category threshold Scheme

Table 27 - Product Category Specific Model Weights

Product Category Specific Model Weights

Criteria	Fashion	Hygiene	Reusables	Home Decor
Recycled Materials in Product	14.03	-	11.05	13.29
Renewable Materials in Product	13.91	11.28	11.04	13.07
Organic Farming	9.32	16.29	-	3.90
Recyclability of Waste	1.68	-	8.14	1.19
Renewable Energy	-	-	4.44	-
Biodegradable	7.89	11.25	10.34	10.18
Product Recyclability	9.82	-	9.36	11.35
Product Repairability	11.45	1.76	6.22	4.40
Product Disassembly	-	0.74	1.60	-
Product Disposability	8.38	3.35	10.62	8.14

Table 28 - Product Category Specific Model Weights (Continued)

Product Category Specific Model Weights				
Product Packaging	2.75	3.36	4.28	1.90
Eco-Label	5.88	18.66	2.17	6.63
ISO 14000 Certification …	-	-	0.52	-
CO_2 Sequestration	7.64	-	3.04	-
Green Delivery	1.34	4.76	3.60	2.57
Ecological Mitigation Effort	-	-	1.00	-
Welfare Impact	-	2.37	6.30	-
Locally Employed	4.33	5.05	2.44	5.05
Animal Welfare	1.12	21.13	-	-
Ratio of Waste Produced	0.46	-	3.84	-
Water Consumed	-	-	-	3.34
Energy Consumed	-	-	-	8.93
Electricity Consumed by Product	-	-	-	6.06

It is important to notify due to the high level of interaction amongst participants, the identification of a suitable panel, both relevant and experienced, is of paramount importance when considering the approach in question. The abovementioned scoring and weighting processes were performed using the software M-MACBETH. The software automates the additive aggregate model calculations in order to derive overall global scores and allows to conduct sensitivity analysis and the robustness analysis.

5 Results

This chapter presents the outcomes of web-Delphi process and the results obtained from the application of the developed multicriteria model in the evaluation of consumable products. The web-based Delphi was applied in accordance to the methodology described in the Section 3 (sub-section 3.4.2). The multicriteria model was implemented as specified in Section 3 (sub-section 3.6).

5.1 Web Delphi Results

The objective of this web-Delphi process was to elicit the expertise of the specialists to identify potential criteria necessary for the construction the multicriteria model. The panel were requested to indicate their agreement or disagreement with the criterion's relevance for assessing sustainability of products. The results are presented in an agglomerated form, concerning the response rate of the panel and on an individual level, where participants' responses on each criterion are portrayed.

5.1.1 Participants Response Rate

In the 1st round of the Delphi survey, a total of 55 participants were invited to engage in the process. From the 55 invited participants, a total of 49 participants responded to the questionnaire contributing to an 89% response rate in corresponding to a dropout rate of 11% among panel members. The 49 respondents from the 1st round, were invited to take part in the 2nd round of the process. Of the 49 invited participants, a total of 44 participants answered the survey resulting in an 90% adherence corresponding to a dropout rate of 10% among the participants. The 44 respondents from the 2nd round, were invited to partake in the 3rd and

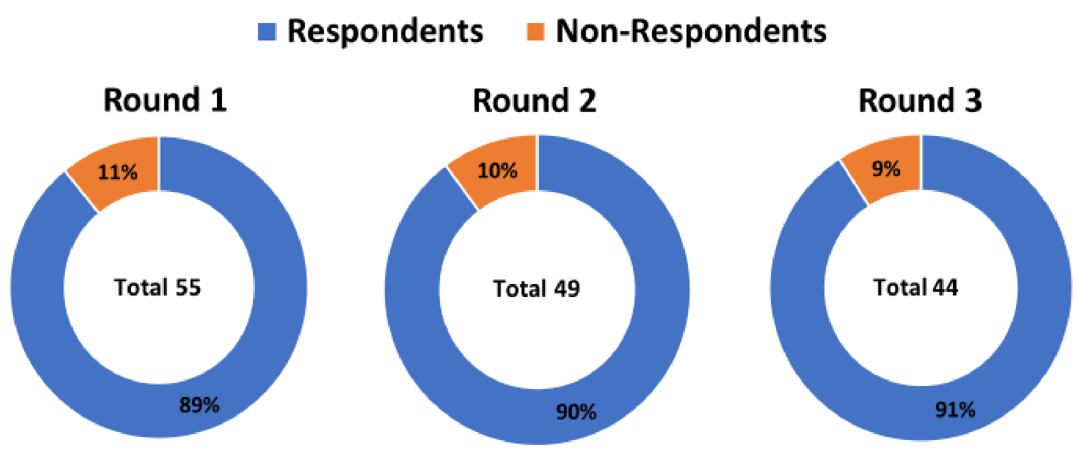


Figure 15 – Web Delphi Response Rate

final round of the Delphi process. In the final round 40 members of the remaining panel replied to the questionnaire constituting a 91% response rate relative to a dropout rate of 9% among the panel. Figure 15 showcases the percentage of respondents and non-respondents across the three rounds.

5.1.2 Criterion Agreement

A total of 34 criteria were proposed to the panel, and amongst them 31 criteria (91.2%) reached group agreement and were approved as relevant to appraise product sustainability. The agreement results in the areas of concern are as follows: Environmental (22), Social (3) and Economical (6). Three indicators presented lack of agreement and were subsequently rejected. The group agreement reached in each criterion was stable across the three rounds. Table 28 displays the indicator approval with respect to areas of concern and dimensions.

Table 29 - Indicator Proposed, Approved and Rejected by Areas of Concern and Dimensions

Area of Concern & Dimension	Proposed	Group Agreement		No Agreement
		Approved	Rejected	
Environmental	23	22	0	1
Raw Materials	5	5	0	0
Manufacturing Phase	2	2	0	0
EcoLabels & Certification	3	3	0	0
Product End of Life	5	4	0	1
Product Packaging	2	2	0	0
Renewable Energy	1	1	0	0
Toxicity	2	2	0	0
CO_2	1	1	0	0
Delivery	1	1	0	0
Mitigation Impact	1	1	0	0
Social	5	3	0	2
Economical	6	6	0	0
Total	34	31	0	3

Figure 16 displays the break-down of indicator approval with respect to areas of concern.

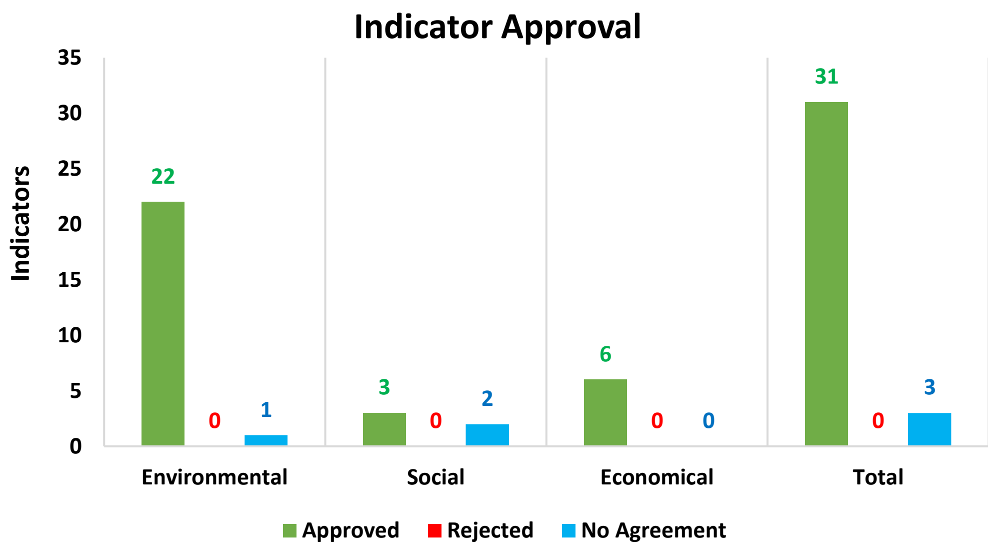


Figure 16 - Indicator Approval with respect to Areas of Concern

Participants had the option to comment on the criteria over the course of 3 rounds. Comments were shared in the 1st and 3rd Rounds. Majority of the comments were generic and supportive of the criteria proposed for sustainability assessment. As stated in one comment, "*Using wood over steel or concrete for construction applications is by far, way more sustainable*". Although concern was raised on the measurability of the indicators, none posted to objection in the selection of the criterion. The panel also commented on the nature recycling processes.

Indicators were approved on the basis of the "Agreement Protocol" depicted in section Figure 4. Out of the 31 approved indicators, 21 were approved by *Absolute Majority* (SA > 50% and SD < 35 %), 10 were approved by *Qualified Majority* (A and SA > 75%). On 3 indicators, *No Agreement* was reached and hence were rejected. The *Qualified Majority* rule resulted in the increase in the number of approved indicators in all areas of concern. A total of 42% of the indicators were approved by this rule. No indicator was rejected on *Absolute Majority* (D and SD > 50%) in the Delphi process, though high amount of disagreement was present under the *Social* area of concern.

Table 29 represents indicator stratification by area of concern and dimensions and the percentage of agreement reached on the indicators approved by group majority according to the agreement protocol.

Table 29 - Indicator Stratification by Agreement Protocol

Area of Concern	AM-A	QM-A	NM	AM-R
Environmental	14	8	1	0
Social	1	2	2	0
Economical	6	0	0	0
Total	21	10	3	0

AM-A, Approved by Absolute Majority (SA = > 50% and SD + D < =35%), QM-A, Approved by Qualified Majority (A + SA = > 75%), NM, No Majority reached and AM-R, Rejected by Absolute Majority (D + SD = > 50%)

After the 3 rounds of the web-Delphi process resulted in a higher convergence of group opinion towards agreement on the indicator's relevance. Table 30 presents the aggregated response on each indicator and the decision based on the agreement protocol, by area of concern and dimension.

Table 30 - Aggregated Responses on each indicator by area of concerns and dimension

Dimension and Indicators	Aggregated Response (%)						Decision
	SD	D	NAD	A	SA	NA	
Environmental							
Raw Materials							
Virgin Material Content	0.0	8.0	15.0	48.0	30.0	0.0	QM-A
Renewable Material in Product	0.0	0.0	0.0	12.5	87.5	0.0	AM-A
Organic Farming	0.0	5.0	10.0	62.5	20.0	2.5	QM-A
Organic Raw Materials	0.0	2.5	5.0	67.5	22.5	2.5	QM-A
Recycled Materials in Product	0.0	0.0	2.5	20.0	77.5	0.0	AM-A
Manufacturing Phase							
Waste Recyclability	0.0	0.0	0.0	22.5	77.5	0.0	AM-A
Recyclability of By-Products and Waste	0.0	0.0	2.5	17.5	80.0	0.0	AM-A

SD, Strongly Disagree, D, Disagree, NAD, Neither Agree nor Disagree, A, Agree, SA, Strongly Agree, NA, No Answer, AM-A, Approved by Absolute Majority (SA = > 50% and SD + D < =35%), QM-A, Approved by Qualified Majority (A + SA = > 75%), NM, No Majority reached

Table 30 - Aggregated Responses on each indicator by area of concerns and dimension (Continued)

Dimension and Indicators	Aggregated Response (%)						Decision
	SD	D	NAD	A	SA	NA	
Environmental							
Raw Materials							
EcoLabels & Certification							
Eco-Label Recognition	0.0	5.0	10.0	50.0	35.0	0.0	QM-A
Types of Eco-Label Accreditation	0.0	5.0	15.0	55.0	22.5	2.5	QM-A
ISO Certification of Intermediary Suppliers	0.0	2.5	12.5	62.5	20.0	2.5	QM-A
Product End of Life							
Product Recyclability	0.0	0.0	0.0	25.0	75.0	0.0	AM-A
Product Repairability	0.0	5.0	7.5	25.0	62.5	0.0	AM-A
Product Disassembly	0.0	7.5	22.5	47.5	22.5	0.0	NM
Ease of Product Disassembly	0.0	7.5	12.5	47.5	32.5	0.0	QM-A
Product Disposability	0.0	2.5	10.0	30.0	55.0	2.5	AM-A
Product Packaging							
Reuse of Packaging	0.0	0.0	2.5	17.5	80.0	0.0	AM-A
Recycling of Product Packaging	0.0	0.0	0.0	30.0	70.0	0.0	AM-A
Renewable Energy							
Renewable Energy Utilization	0.0	0.0	10.0	7.5	80.0	2.5	AM-A
Toxicity							
Toxicity in Product Manufacture	0.0	5.0	0.0	22.5	72.5	0.0	AM-A
Toxicity of Waste	0.0	5.0	0.0	25.0	70.0	0.0	AM-A
CO_2							
CO_2 Sequestration	0.0	2.5	7.5	30.0	60.0	0.0	AM-A
Delivery							
Green Delivery	0.0	0.0	5.0	27.5	65.0	2.5	AM-A
Mitigation Impact							
Ecological Mitigating Effort	0.0	2.5	7.5	35.0	47.5	7.5	QM-A

SD, Strongly Disagree, D, Disagree, NAD, Neither Agree nor Disagree, A, Agree, SA, Strongly Agree, NA, No Answer, AM-A, Approved by Absolute Majority (SA = > 50% and SD + D < =35%), QM-A, Approved by Qualified Majority (A + SA = > 75%), NM, No Majority reached

Table 30 - Aggregated Responses on each indicator by area of concerns and dimension (Continued)

Dimension and Indicators	Aggregated Response (%)						Decision
	SD	D	NAD	A	SA	NA	
Social							
Welfare Impact	0.0	2.5	15.0	42.5	35.0	5.0	QM-A
Locally Sourced	0.0	2.5	7.5	30.0	60.0	0.0	AM-A
Culture Preservation	0.0	15.0	12.5	52.5	20.0	0.0	NM
Number of Employees	5.0	32.5	40.0	10.0	12.5	0.0	NM
Animal Welfare	0.0	2.5	20.0	32.5	42.5	2.5	QM-A
Economical							
Ratio of Waste Produced (%)	0.0	0.0	5.0	20.0	75.0	0.0	AM-A
Ratio of Raw Material Consumed (%)	0.0	0.0	5.0	40.0	55.0	0.0	AM-A
Electricity Consumed	0.0	0.0	5.0	27.5	67.5	0.0	AM-A
Water Consumed	0.0	0.0	0.0	27.5	72.5	0.0	AM-A
Energy Consumed	0.0	2.5	7.5	30.0	60.0	0.0	AM-A
Electricity Consumed	0.0	0.0	15.0	35.0	50.0	0.0	AM-A

SD, Strongly Disagree, D, Disagree, NAD, Neither Agree nor Disagree, A, Agree, SA, Strongly Agree, NA, No Answer, AM-A, Approved by Absolute Majority (SA = > 50% and SD + D < =35%), QM-A, Approved by Qualified Majority (A + SA = > 75%), NM, No Majority reached

The web-Delphi process was successful in procuring agreement on majority of criteria for assessing sustainability in products. In this process, a multidisciplinary panel (6 stakeholder and 49 experts) was consulted in three rounds, from which 31 were approved out of a list of potential 34 indicators and 3 indicators were rejected as they did not achieve the required group agreement in order to be selected. The panel reached agreement on indicators from multiple areas of concern and dimensions, reinforcing the multidimensional and holistic nature of sustainability.

5.2 Evaluation Model Results

The results of the web-Delphi process were wielded to be fed in the construction process of the evaluation model. These results correspond to the score achieved by the alternatives on the product assessment tool with regards to their sustainability. In order to demonstrate the multiple potential applications of the model, products from a diverse set of categories were considered. The models created for the specific categories are as following: Reusables, Fashion, Hygiene and Home Decor. The products were evaluated on the "Reusables" model and the results were analyzed. Only products pertinent to this category were assessed. The same procedure was carried out for other categories and their results are represented in Appendix D. In this case study, products from the PLANETIERS' platform were selected for appraisal.

5.2.1 Reusable Model Results

The evaluation of the products was carried out first on the "Reusables" model. This model was useful to analyze the broad spectrum of products in comparison to each other across the Reusables product categories. The results are displayed in Figure 17, in the form of an "Overall Thermometer Scale", depicting standings obtained by the products on the Reusables sustainability assessment tool. The reference level "Good" represents the benchmark score "100" where a hypothetical alternative has a performance level affiliated with *Good* reference level in all criteria. Likewise, the reference level "Neutral" represents the benchmark score "0" where a hypothetical alternative has a performance level affiliated with *Neutral* reference level in all criteria. The model proposed the product "*Plate*" as the most sustainable option in commensurate with the decision makers judgments. According to the category thresholds defined previously (sub-section 4.2.6), *Plate* has scored an overall score of "111.02" is approved and considered sustainable. The product "*Bag 1*" scored an overall score of "103.08" and is considered sustainable and approved. The overall scale represents that the remaining products are less sustainable in comparison. The items labelled as "*Bottle 2*" and "*Bag 2*" have scored in the negative, "-6.32" and "-8.22" respectively. Their scores represent that the products were deemed

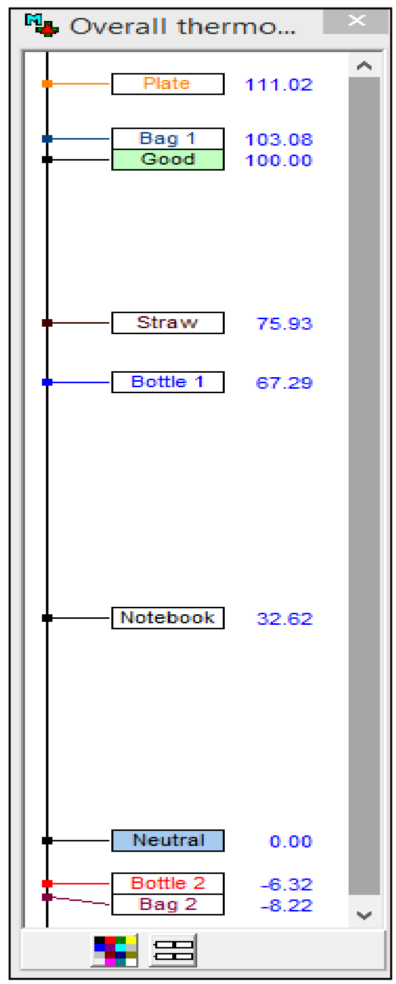


Figure 17 - Reusable Model Results

non-sustainable and henceforth rejected. The products, “*Straw*” and “*Bottle 1*” scored in the range of 50 and 100. These products will be regarded sustainable and approved but will be monitored, as defined by the category thresholds. The product “*Notebook*” achieved an overall score of “32.62”. The performance levels of *Notebook* were analyzed on the criteria “Social Welfare”. As the performance level of item *Notebook* was equivalent to the “Good” reference level in criterion Social Welfare, it was regarded sustainable and approved.

The differences profiles graph shown in Figure 18, compares the products “Plate” and the second highest scoring option “Bag 1”. The weighted differences profile allows to analyze the extent to which the differences compensate in favor of the first or the second option. In majority of the criteria they have scored equally hence the presence of null difference between them and is depicted by “0.00”. The “Green” weighted bars under *Product Disposability* showcases the Plate obtained a difference of “+5.31” in favor.

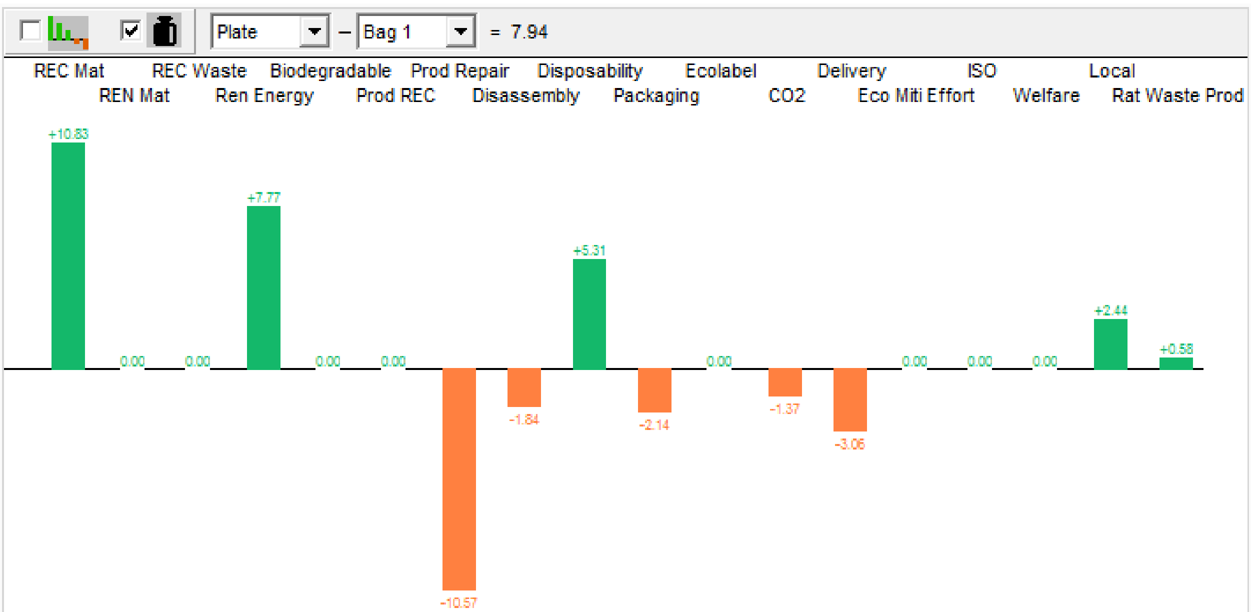


Figure 18 - Weighted Differences Profiles, "Plate vs Bag 1"

In the criterion *Product Repair*, Bag 1 obtained a difference of “-10.57” in favor. The sum of the weighted differences of scores in the criteria favorable to Plate compensates more than the sum of the weighted differences of scores in the criteria favorable to Bag 1. The overall difference between the two products is “7.94”.

When we compare the weighted differences in profiles of Straw and Bottle 1, the overall difference between them is of “8.64” (depicted in Figure 19). Although Bottle 1 proved to be better in greater number of criteria than Straw, the weighted differences compensate in favor of Straw.

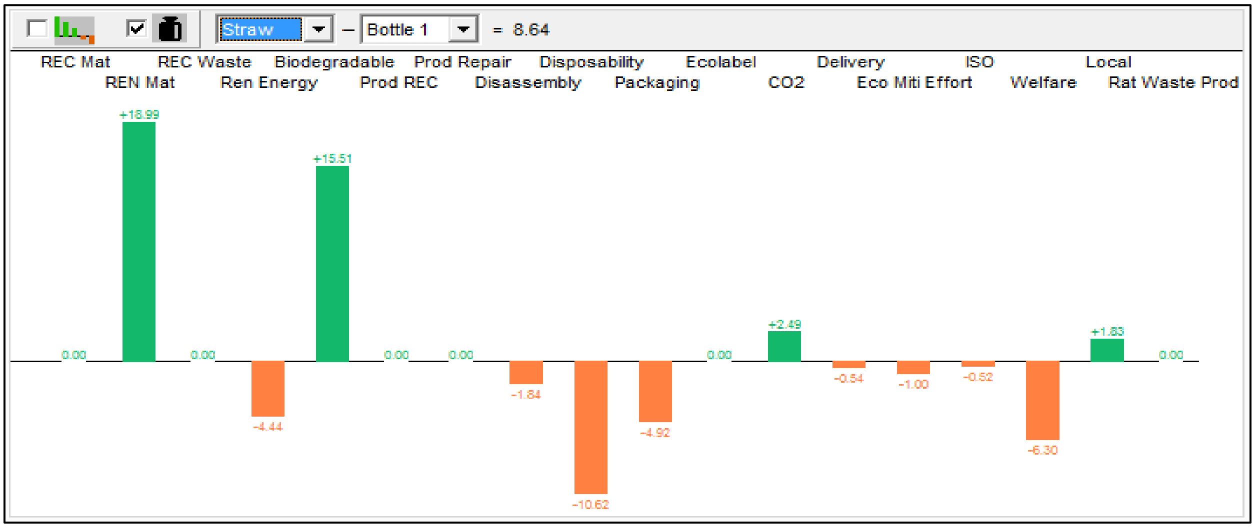


Figure 19- Weighted Differences profiles, "Straw vs Bottle 1"

The products score on two criteria, Product Repair and Welfare Impact, are compared in Figure 20 on a two-dimensional “XY Map”. Even though Plate does not fall on the efficient frontier as compared to the other products, and has scored negative in Product Repair, it still manages to achieve the highest overall score.

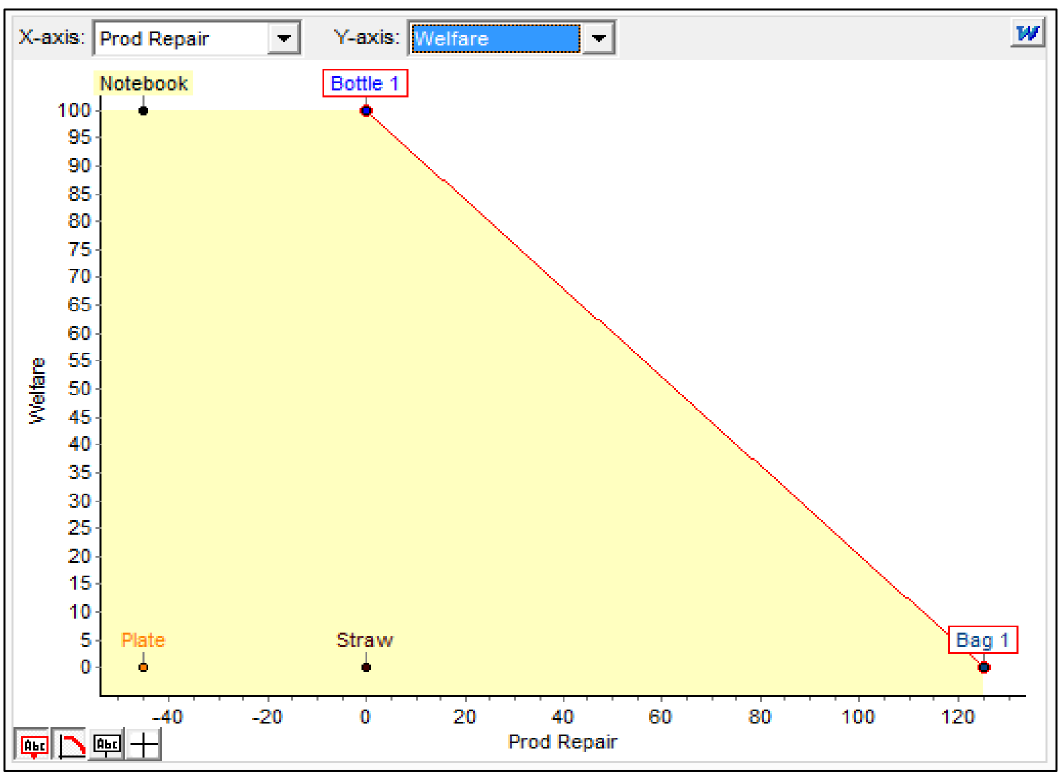


Figure 20 - XY Map, "Product Repair vs Welfare Impact"

(Phillips, 1984) states “a model is considered requisite if it is sufficient in form and content to resolve the issues at hand”. The reliability of the model is ascertained by the recursive processes involved in the model building activity as indicated by the reverse arrows in Figure 1. When evaluating alternatives, it is also necessary to perform a critical analysis on the validity of the model. The resilience of an option is assessed by conducting sensitivity and robustness analysis when alterations within the model are induced. Sensitivity and robustness analysis aid in determining the impact of modifying the parameters of the model on the resulting outcomes. This supports the decision maker in identifying potential aspects of the model that need further adjustments. These adjustments can range from tuning of the model parameters to the restructuring of the model itself.

The stability of the additive value model was determined by analyzing small variations in the weight of “Welfare Impact” criterion. The sensitivity analysis graph shows the variations in the overall score of the options when the criterion weight varies. The sensitivity analysis graph in Figure 21 shows that Plate remains the most sustainable product even when the weight of *Welfare Impact* is increased from 6.30% to nearly 34.8%. After the weight has increased beyond that, the product Notebook becomes the most sustainable in comparison.

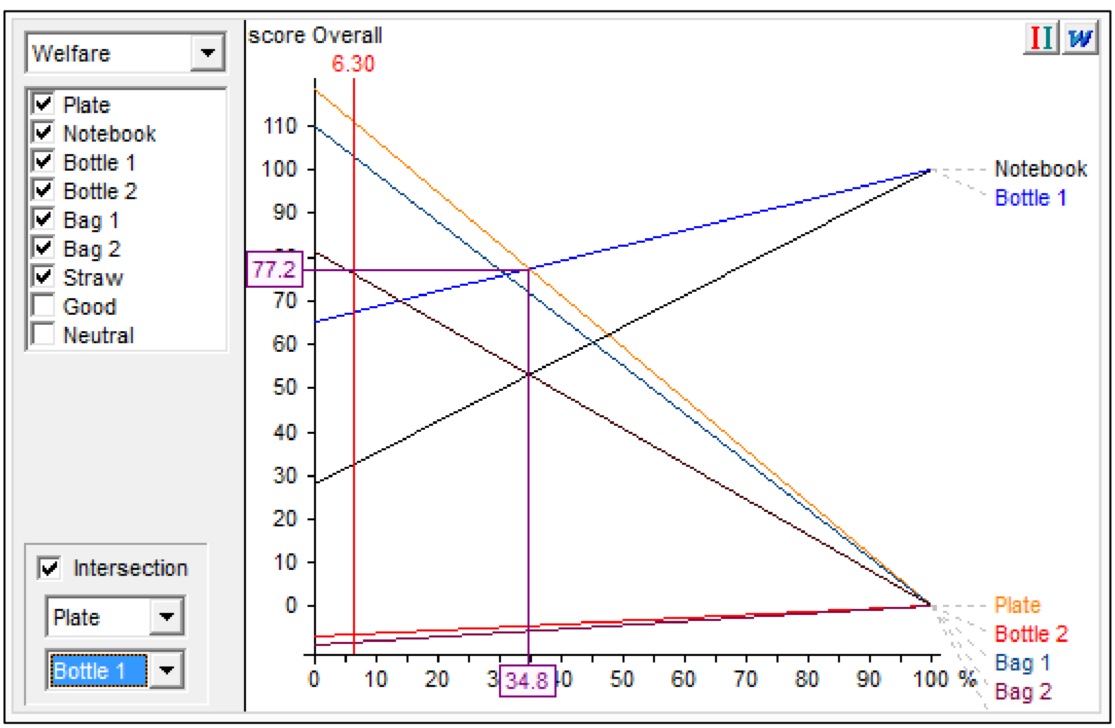


Figure 21 - Sensitivity Analysis "Welfare Impact"

Decisions often involves incomplete or uncertain information and it can be useful to determine conclusions from varying amounts imprecision in information. (Bana e Costa *et al*., 2012) mentions conclusions can be drawn from only ordinal or/and pre-cardinal intra-criteria and inter-criteria information. M-MACBETH software provides the robustness analysis function for this purpose. In the absence of Cardinal Information

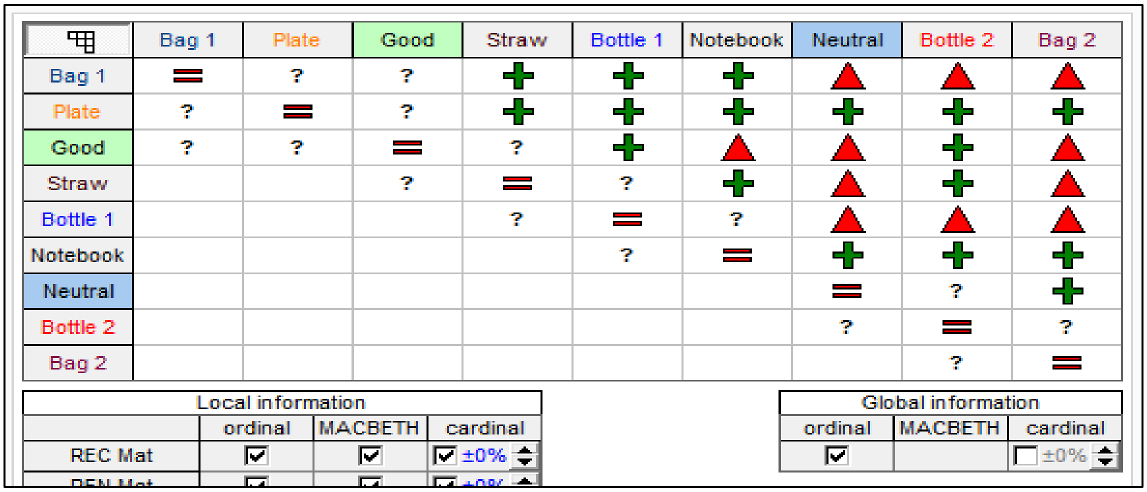


Figure 22 - Robustness Analysis in absence of Cardinal Global Information

in the Global Information section, as shown in Figure 22, it cannot be determined if “Plate” or “Bag 1” are the highest-ranking products. Once the cardinal information is added on the global information section, as depicted in Figure 23, the ranking of the products becomes evident.

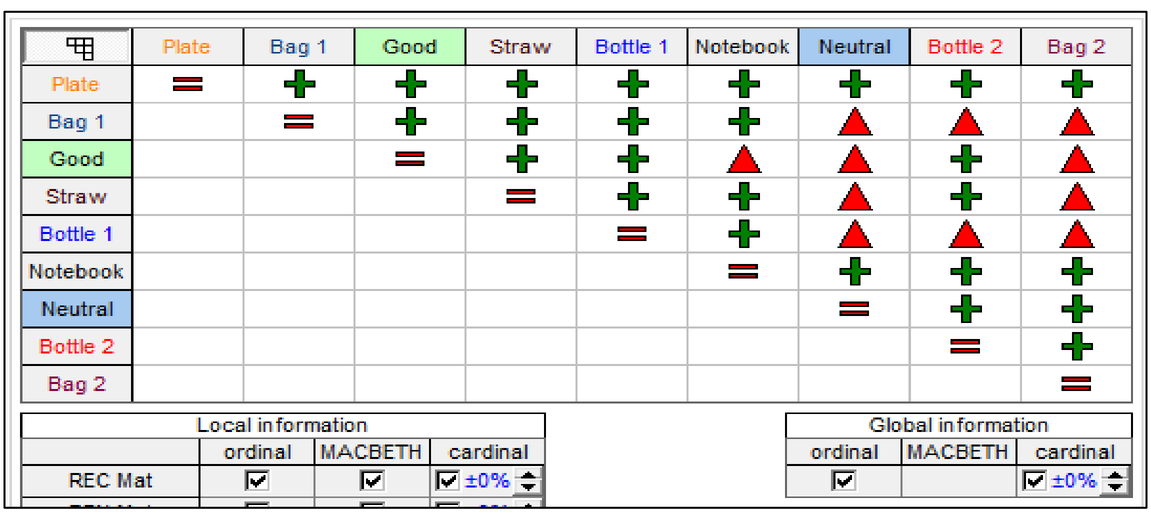


Figure 23 - Robustness Analysis with Cardinal Global Information

This function is particularly helpful when uncertainty is present in the information. In the case when an uncertainty of 12%, 11% and 11% is introduced in the criteria “Renewable Materials in Product”, “Biodegradable” and “Green Delivery”, depicted in Figure 24, Plate remains the most sustainable product, however, not information is available to conclude that if Bag 1 is globally more attractive than the reference level “Good”.

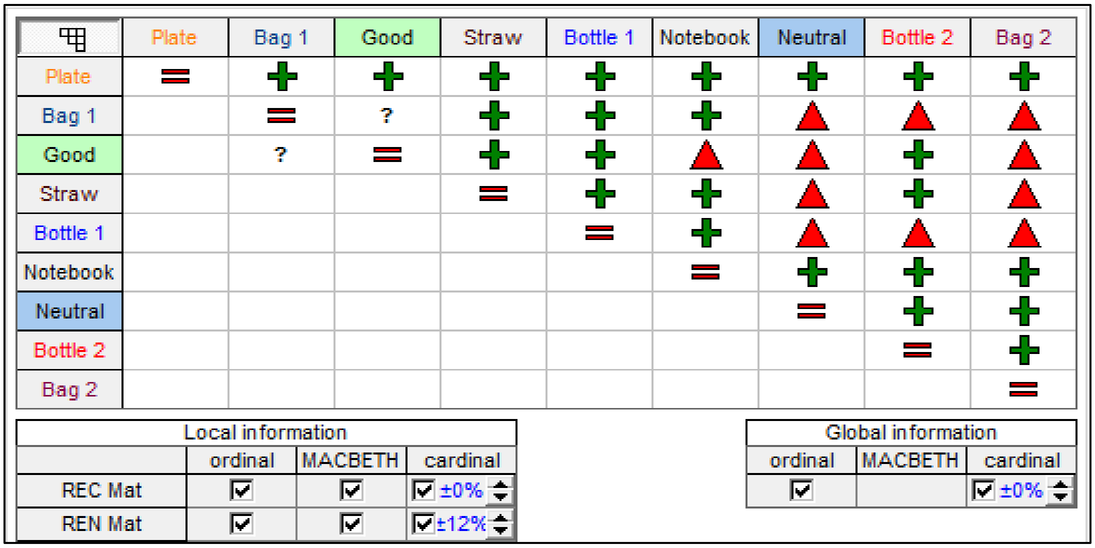


Figure 24 - Robustness Analysis with Uncertainty

The establishment of product category specific assessment tools allowed for more reliable and robust evaluation of products, as criteria that were deemed unnecessary for certain categories were removed. This also facilitated in assigning higher weightages to certain criteria in one model and a lower weightage in another model in consideration of the category assessed. This capability was not present in the generic model which resulted in “lack of trust” with the results. Given the broad spectrum of products available, the construction of specific models was considered appropriate and the results obtained also corresponded with the expertise knowledge of the decision makers.

One of the outcomes of the methodological framework was the creation of a “Supplier Product Information” form. The descriptors of each criterion were used to design the form. The purpose of this form will be to acquire product information to facilitate product appraisal. This proved to be an important component as this form greatly enhanced the efficiency and speed of the evaluation process. It formalized a standard set of information necessary for product appraisal further streamlining the process.

6 Discussion

The application of the framework resulted in the employment of a web-based Delphi process and construction of a multicriteria model utilizing MACBETH approach to develop a tool for assessing sustainability of products. The tool was tested on the PLANETIERS' products for this case study. The application of Delphi methodology and proposed MCDA approach of this case study are discussed in this section. This chapter is divided into two sections; 1) Delphi Survey and 2) Multicriteria Model.

6.1 Delphi Survey

The aim of this study was to develop a sustainability assessment tool for evaluating household commodities and consumer products through an employment of a web-Delphi process designed to collect expert opinion of the participants.

6.1.1 Welphi Contribution

The use of the Welphi platform to deliver the questionnaires greatly enhanced the speed of the Delphi process as the need for face to face interaction was eliminated hence saving time and resources. The automated intrinsic procedures such as data entry, responses and analysis enabled transitioning towards the subsequent rounds much more swift and easier. The Welphi platform enabled monitoring of participant responses which was crucial to send timely reminders and reduce panel drop-out.

6.1.2 Delphi Response

A high response rate of 89%, 90% and 91% obtained during the first, second and third rounds, showcased that the respondents could see the relevance and purpose of the Delphi process. This level of response exceeded expectations as the panel being a large and diverse whilst also comprising of multiple nationalities. Although the retention rate was considerably high in this particular case study, there are several reasons for dropping out of the process such as lack of time, uncertainty, deviation from the group opinion, among others. The questionnaire was only distributed to those participants who had already confirmed their availability beforehand and personal communication methods were employed to reach out to panelists resulting in reduced dropout rate as confirmed by (Belton *et al*., 2019).

6.1.3 Panel

By requesting level of agreement, there was no need to force a consensus among the opinions of the experts. Group feedback was given in percentage and absolute form which was easy in comprehension for the participants. The validity of the results was confirmed by the comments of the participants as the participants voiced in support of the criteria scrutinized. Although a high percentage of agreement was reached, through Absolute or Qualified Majority, the heterogenous panel was not represented equally and could have potentiated a differentiation in the results obtained if equal representation was present.

6.2 Multicriteria Model

It was necessary to establish the type of approach implemented in the construction of the evaluation model once the criteria had been identified. Considering the number of criteria and the multiple areas of concern, the MACBETH approach was selected, as it has been used in conjunction with Delphi process.

6.2.1 MACBETH Approach

The use of MACBETH methodology proved to be extremely suitable for the construction of the multicriteria model. The benefit of using of MACBETH was by far the reciprocal nature that allowed continuous feedback of the stakeholders during the development process and the capability of constructing multiple criteria under tailor-made situations. The recursive model development procedure under MACBETH's non-numerical approach offers an advantage of interactive verification of the reliability of the preference information elicited (Bana e Costa *et al*., 2012). The use of this method in eliciting qualitative answers was extremely helpful as it relieved the technical burden placed upon the decision makers. The more fluent decisionmakers find the non-numerical elicitation technique easier and more satisfying, (Fasolo and Bana e Costa, 2014), and this was reflected by the continued interest on part of the decision makers during the decision conferences.

6.2.2 Structuring Value Tree

The complexity of the value tree is mostly as a result of the ambiguity and broad definition of sustainability leading to a large number of criteria. Given the nature of the criteria identified and further considering their number, it was decided that a bottom-up approach to be the most suitable in the model building process of the product assessment tool. Each evaluation criteria were operationalized by an associated descriptor of performance which represented significant difficulties, as it was extremely difficult to establish substantial cause-effect relationships especially concerning the criteria under the "Economical" branch of the value tree. The established criteria consisted of almost qualitative or constructed descriptors rather than quantitative due to the lack of available data and measurement capabilities on the side of the suppliers.

6.2.3 Generation of Value Function

The structuring and evaluation components of the multicriteria framework required constant adjustments and refinements to accurately depict the judgements of the decision makers thus gaining verification and acceptance of the actors involved. The use of same value functions from the generic to category-based models was validated by the decision makers yet the most preferable scenario would have been to create value functions keeping in consideration the category of products being assessed.

6.2.4 Determination of Criteria Weights

The use of “Good” and “Neutral” reference levels in each criterion enabled the development of a weighting process not dependent on fixed extreme references. As explained by (Belton and Stewart, 2002), “The use of central rather than extreme reference points may guard against inaccuracies arising because of possible non-linearity in values occurring at extreme points, a factor which is particularly important in the assessment of weights.” This procedure of utilizing “Good” and “Neutral” reference levels allows construction of scaling constants in situations where the weights must be determined before the assessment of options (Bana e Costa *et al*., 2002).

In certain criteria, only two performance levels could be attributed hence the descriptors acted as extreme fixed levels in generation of weights. In other criteria one of the two “Neutral” or “Good” reference levels were at the extreme levels thus curtailing the score of alternatives to not exceed beyond these levels. The calculation of weighting coefficients proved to be quite problematic and time consuming considering the number of evaluation criteria involved resulting in a large size of the judgment matrix and the presence of scaling constants at different hierarchical levels of the value tree. The number of criteria also limited the weighting coefficients assigned to be of similar in value in the generic model. This was resolved when category-based models were constructed.

6.2.5 Supplier Form and Product Classification

One of model outputs was the design of a “Supplier Product Form” with an aim to elicit product specifications from the supplier for product assessment. This proved to be an important component as this form greatly enhanced the efficiency and speed of the evaluation process. It formalized a standard set of information necessary for product appraisal further streamlining the process. The form based on the constructed descriptors of performance automize the process as all the requisite information was now available from the suppliers.

The overall score resulted from the additive value model facilitated in sorting and cataloguing the products under the category thresholds for product classification. The thresholds introduced were simple to comprehend and implement, the products scoring within the threshold range created complications for the decision makers.

6.2.6 Generic and Category Specific Model

It was deemed necessary to construct a generic model and category specific models as it showcased the flexibility and adaptability of the model with regards to different scenarios. The generic model can be used to evaluate all product types, ranking them and facilitate comparison of a diverse set of products on a single evaluation model. However, this would constitute a problem, as several criteria were affecting the result where their application was not relevant to the category the product belonged to. The generation of multiple category specific models provided a much more accurate depiction of results and the sustainability score attained and eliminated irrelevant criteria for that category. This also facilitated in assigning higher weightages to certain criteria in one model and a lower weightage in another model in consideration of the category assessed. This capability was not present in the generic model which resulted in "lack of trust" with the results. The categories Reusables, Hygiene, Fashion and Home Décor were chosen as they comprise of items most commonly traded.

6.2.7 Decision Support System

M-MACBETH offers tools to easily build a value tree, construct descriptors, show value functions, assign weights, rank alternatives, perform sensitivity and robustness analysis. The use of this software as a decision aid to the MACBETH approach was extremely helpful. The display of the range of value scores on the thermometer scale served as a visual aid in comprehending the relative attractiveness of the judgements. The systematic inconsistency checks eliminated errors occurred during the structuring phase and enabled the opportunity to reconsider the made judgements. Multiple ways of fixing the inconsistencies pointed out by the software (when more than one was present), forced a more thorough and complete building process.

6.2.8 Limitations

The interpretation of the criteria reveals many constructed descriptors have been assigned in all three areas of concern. This in fact due to inability on the part of the product manufacturer to measure the criteria in quantitative terms. Not having continuous quantitative descriptors resulted in assigning wide range descriptors to measure a numeric feature. This greatly inhibits the desired accuracy in assessing sustainability.

The model also exhibits lack of economic impact on the manufacturer resulting from the production of the product in terms of energy or water utilized in the production process. This is in part due to products primarily manufactured by small and medium enterprises which do not have the instrumentation to measure these impacts. Another reason for their exclusion is that each product has a unique consumption of utilities consumed hence constructing a measurable descriptor that would cover a wide range of products would be extremely difficult.

The ambiguous nature of sustainability and the restrictions posed by the scope of the problem context, resulted in construction of very few descriptors in certain criteria. This led to assigning "Good" and "Neutral" reference levels to one or both descriptors. The value functions generated were limited by the reference level scores. This culminated in the products achieving a global score in close proximity to “100” although they were considered sustainable.

6.2.9 Contribution to Literature

This study provides a comprehensive analysis of the application of a Delphi process for the construction of the model. The thesis adds to literature from a 1) practical aspect; by establishing a consensus among a diverse and geographically dispersed group of experts and stakeholders on which indicators are mostly relevant for appraising sustainability of products, from a 2) methodological standpoint, by proposing participatory methods that can be used in the field of sustainability assessment and from a 3) technological viewpoint, by presenting a web-platform that enables the use of participatory processes.

The use of MACBETH approach, founded on measuring difference in attractiveness based on qualitative pairwise comparisons, combined both social and technical aspects of the model building process. This study contributes to the literature by exploring the use of MACBETH methodology in assessing sustainability and by presenting “M-MACBETH”, a decision support software, as a means to aid the use of participatory processes.

6.2.10 Possible Applications and Future Work

Minimal literature was present on assessing products for sustainability and certainly no tool had been developed for such purposes. This in itself turned out to be a novel concept which has great potential due to the growing awareness of sustainability in the society and the ever-increasing demand of sustainable brands and products. The product assessment tool now serves to methodize the product evaluation process which was previously being carried out in an ad hoc manner and bringing about coherency and consistency in the assessment method. The tool also relieves the workload as the generated supplier form automizes the task of eliciting product performance levels hence accelerating the evaluation process.

The derivation of specific models from a generic model further augmented the application capability of the model. Given the fact that it can be tailored for category specific products, it can be applied for a broad spectrum of products further demonstrating adaptability to different situations. New marketplaces have emerged trading in and connecting consumers to sustainable merchandise, hence, the possibilities of application of this model are considerably extensive.

The framework, though designed for assessing sustainability of consumer products, can be implemented in numerous scenarios within the field of sustainability and external domains. The framework inculcates all the necessary steps for the construction of a model in order to assess alternatives pertaining to countless specializations hence augmenting the usability of the framework.

7 Conclusions and Future Work

The increase in the awareness of resource depletion and deteriorating ecosystem has led to a significant effort both from the public and private sector to combat municipal waste generation. The demand for sustainable and ecofriendly items and products has drastically increased and many corporations, including PLANETIERS, have stepped in to bridge the gap. The non availability of a recognized standard results in ad hoc evaluations which do not incorporate all potential criteria for examining the product. PLANETIERS' wished to implement a system that would introduce consistency in their appraisal procedure and automize the process. The work performed in thesis aimed to develop a methodological framework with a purpose of creating a product assessment tool to assess the sustainability of the products.

The first task concerned with the development of a suitable tool that could be employed in consideration of the problem context. A literature review was conducted in identifying methodologies and techniques to develop a product assessment tool. However, it yielded no results as such a tool had not been developed. Hence a new methodological framework needed to be developed for the construction of a multicriteria model for the creation of the assessment tool. A detailed literature review was carried out to identify methods and approaches that would constitute the model building activities of the framework, as presented in Chapter 2, and the MACBETH approach was selected to be the most pertinent for this case study. The initial phase of the framework required identification of potential parameters and criteria necessary for assessing sustainability pertaining to commodities and consumables. This required the application of a participatory technique to elicit the opinions of experts with regards to relevance of proposed criteria in assessing sustainability. Delphi process was chosen with an aim at collection of judgments from an enlarged group of participants in order to posteriorly feed the model creation phase. After identification of the relevant criteria, MACBETH approach was utilized to build the evaluation model. The application of the MACBETH methodology was carried out by means of a decision conference, a socio-technical process, that enabled the interaction between the decision makers and the facilitator. From this decision conference resulted in a defined evaluation model, consisting of a value tree, descriptors of performance, value functions and weighting coefficients. The construction of the model was aided by M-MACBETH, a decision support system, tailored for the MACBETH approach. The developed tool was implemented to evaluate the sustainability of the products on the PLANETIERS' database.

The results of the methodological framework were presented in Chapter 5 and was divided into two sections; first pertaining to the Delphi process and the second corresponding to the product assessment via the established assessment tool. This study provides a comprehensive and sound analysis of the application of a Delphi process for the construction of the model. With regards to the Delphi process, the thesis adds to literature from a 1) practical aspect; by establishing a consensus among a diverse and geographically dispersed group of experts and stakeholders on which criteria are mostly relevant for appraising sustainability of products, from a 2) methodological standpoint, by proposing participatory methods that can be used in the field of sustainability assessment and from a 3) technological viewpoint, by presenting a web-platform that enables the use of participatory processes. Delphi results proved to be very useful in the model's construction. The use of MACBETH approach, founded on measuring difference in attractiveness based on qualitative pairwise comparisons, combined both social and technical aspects of the model building process. This study contributes to the literature by exploring the use of MACBETH methodology in assessing sustainability and by presenting "M-MACBETH", a decision support software, as a means to aid the use of participatory processes. The MACBETH approach proved to be "C*lient Friendly*" for decision makers in expressing their judgements. The continuous consistency checks and visual aids present in the M-MACBETH software enhanced the model building process. The resilience of the options was tested by conducting sensitivity and robustness analysis and it confirmed the model could be considered requisite. This was ascertained by the recursive development process of the model building activity. The additional output of the "Supplier Product Form" standardized the information elicitation procedure from the suppliers and greatly enhanced the efficiency and speed of the process.

Issues of concern were raised in the model construction process, most notable being the operationalization of the evaluation criteria as the associated descriptors of performance represented significant difficulties in establishing substantial cause-effect relationships. Majority of the established criteria consisted of constructed descriptors, where quantitative would have been preferable, due to the lack of available data and measurement capabilities. The reduced number of associated descriptors, due to the nature of sustainability and problem context, and the use of "Good" and "Neutral" reference levels, curtailed the score of alternatives within the threshold levels. As this study being a development of a novel framework and assessment tool, the methodology may be subject of further improvement, nevertheless, it can be considered that the objective of this thesis has been accomplished.

The tool is currently designed for four product categories and needs to be expanded to the rest of the categories. This would involve analyzing the generic model and selecting criteria pertaining to the chosen category and removing irrelevant criteria. This would allow check for homogeneity within the results. As mentioned previously, the tool adopts major limitations especially regards to many constructed descriptors especially incorporating economic criteria and the hindrance in measuring these criteria. In-depth and detail research needs to be made to identify means of measuring these criteria.

In final conclusion, this study presented, a real-world application, in Sustainability Assessment, a practical assessment tool for evaluating consumer products. This study highlighted the usefulness of relevant experts and stakeholder involvement in the criteria selection process and the multicriteria model building process. The findings from the study can inform future research on sustainability assessment and product appraisal of the importance of accounting multiple points of view. The developed framework can be tailored to not only different aspects of sustainability but also products from different industries. This study contributes to the literature by proposing Delphi and MCDA approaches alongside employment of web-based platform and decision support system can be used in the field of sustainability assessment.

References

GFN. (2019). Website of Global Footprint Network, https://www.footprintnetwork.org/our-work/ecological-footprint/#worldfootprint, accessed July 15th.

NRDC. (2019). Website of Natural Resources Defence Council, https://www.nrdc.org/stories/global-warming-101#causes, accessed on July 15th.

Haddock-Fraser, J. E., and Tourelle, M. (2010). Corporate Motivations for Environmental Sustainable Development: Exploring the Role of Consumers in Stakeholder Engagement. *Business Strategy and the Environment*, 19(8), 527-542. https://doi.org/10.1002/bse.663

Chapman, A., and Shigetomi, Y. (2018). Developing National Frameworks for Inclusive Sustainable Development Incorporating Lifestyle Factor Importance. *Journal of Cleaner Production*, 200, 39-47. https://doi.org/10.1016/j.jclepro.2018.07.302

Dogl, C., and Behnam, M. (2014). Environmentally Sustainable Development through Stakeholder Engagement in Developed and Emerging Countries. *Business Strategy and the Environment*, 24(6), 583-600. https://doi.org/10.1002/bse.1839

OECD. (2017), Tackling Environmental Problems with the Help of Behavioral Insights, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264273887-en

European Commission. (2014). Website of European Commission, https://ec.europa.eu/clima/citizens/eu_en, accessed on April 15th.

PLANETIERS. (2019). Website of PLANETIERS, https://planetiers.com/, accessed on April 20th.

DESA. (2017). United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP/248.

WB. (2019). Website of The World Bank, http://datatopics.worldbank.org/what-a-waste/trends_in_solid_waste_management.html, accessed on April 10th.

NMCE. (1994). Oslo Roundtable on Sustainable Production and Consumption. Norwegian Ministry of Climate and Environment.

EPA. (2016). Website of United States Environmental Protection Agency, https://www.epa.gov/sustainability/learn-about-sustainability#what, accessed on April 1st.

Environmental Science. (2019). Website of Environmental Science Education and Careers, https://www.environmentalscience.org/sustainability, accessed on April 1st.

Brundtland. (1987). Our Common Future. *Report of the World Commission on Environment and Development, United Nations*. Oxford University Press, Oxford.

IUCN/UNEP/WWF. (1991). *Caring for the Earth: A Strategy for Sustainable Living*. Gland, Switzerland.

Shaker, R. R. (2015). The spatial distribution of development in Europe and its underlying sustainability correlation. *Applied Geography*, 63, 304-314. doi:10.1016/j.apgeog.2015.07.009

Munasinghe, M. (2016). Making Development More Sustainable: Sustainomics Framework and Practical Applications (3rd ed.). Colombo, Sri Lanka: Munasinghe Institute for Development (MIND) Press.

NUA. (2016). Website of New Urban Agenda (Habitat III), http://habitat3.org/the-new-urban-agenda/, accessed on April 2nd.

UNFCCC, (2015). Website of United Nations framework Convention on Climate Change, https://unfccc.int/process-and-meetings/the-paris-agreement/what-is-the-paris-agreement, accessed on April 10th.

Burgass, M. J., Halpern, B. S., Nicholson, E., & Milner-Gulland, E.J. (2017). Navigating Uncertainty in Environmental Composite Indicators, *Ecological Indicators*, 75, 268-278.

European Commission – Joint Research Center. (2012). Dealing with Uncertainty in Sustainability Assessment. Luxembourg: Publications Office of the European Union.

Rametsteiner, E., Pulzl, H., Alkan-Olsson, J., & Frederiksen, P. (2011). Sustainability Indicator Development – Science or Political Negotiation. *Ecological Indicators*, 11, 61-70.

Srinivasan, R. S., Braham, W. W., Campbell, D. E., & Curcija, D. C. (2011). Sustainability Assessment Frameworks, Evaluation Tools and Metrics for Buildings and Its Environment – A Review. Proceedings of *Building Simulation 2011: 12th Conference of International Building Performance Simulation Association*, Sydney, 350-356.

Doualle, B., Medini, K., Boucher, X., & Laforest, V. (2015). Investigating Sustainability Assessment Methods of Product-Service Systems. *Procedia CIRP*, 30, 161-166. https://doi.org/10.1016/j.procir.2015.03.008

Sala, S., Ciuffo, B., & Nijkamp, P. (2015). A Systemic Framework for Sustainability Assessment. *Ecological Economics*, 119, 314-325. http://dx.doi.org/10.1016/j.ecolecon.2015.09.015

Devuyst, D. (2001). How Green Is the City? Sustainability Assessment and the Management of Urban Environments. New York City, NY: Columbia University Press.

Finnveden, G., Nilsson, M., Johansson, J., Persson, A., Moberg, A., & Carlsson, T. (2003). Strategic Environmental Assessment Methodologies – Applications within the Energy Sector. *Environmental impact Assessment Review*, 23(1), 91-123. https://doi.org/10.1016/S0195-9255(02)00089-6

Belfield, C., and Levin, H. M. (2010). Cost–Benefit Analysis and Cost–Effectiveness Analysis. *International Encyclopedia of Education*, 3, 199-203. https://doi.org/10.1016/B978-0-08-044894-7.01245-8

Gluch, P., and Baumann, H. (2004). The Life Cycle Costing (LCC) Approach: A Conceptual Discussion of its Usefulness for Environmental Decision Making. *Building and Environment*, 39, 571-580. https://doi.org/10.1016/j.buildenv.2003.10.008

Dodgson, J. S., Spackman, M., Pearman, A. D., & Phillips, L. D. (2009). Multi-Criteria Analysis: A Manual. London: Department for Communities and Local Government.

Keeney, R. L., and Raiffa, H. (1993). Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Canada: Cambridge University Press.

Bana e Costa, C. A., and Vansnick. (2000). Cardinal Value Measurement with Macbeth. In Decision Making: Recent Developments and Worldwide Applications. Boston, MA: Springer, pp. 317-329.

Phillips, L. D., and Bana e Costa, C. A. (2007). Transparent Prioritisation, Budgeting and Resource Allocation with Multi-Criteria Decision Analysis and Decision Conferencing. *Annals of Operations Research*, 154(1), 51-68. https://doi.org/10.1007/s10479-007-0183-3

Cinelli, M., Coles, S. R., & Kirwan, K. (2014). Analysis of the Potentials of Multi Criteria Decision Analysis Methods to Conduct Sustainability Assessment. *Ecological Indicators*, 46, 138-148. https://doi.org/10.1016/j.ecolind.2014.06.011

Klein, S. J. W., and Whalley, S. (2015). Comparing the Sustainability of U.S. Electricity Options through Multi-Criteria Decision Analysis. *Energy Policy*, 79, 127-149. https://doi.org/10.1016/j.enpol.2015.01.007

Roinioti, A., and Koroneos, C. (2019). Integrated Life Cycle Sustainability Assessment of the Greek Interconnected Electricity System. *Sustainable Energy Technologies and Assessments*, 32, 29-46.

Myllyviita, T., Lesknen, P., Lahtinen, K., Pasanen, K., Sironen, S., Kahkonen, T., & Sikanen, L. (2013). Sustainability Assessment of Wood-Based Bioenergy – A Methodological Framework and a Case-Study. *Biomass and Bioenergy*, 59, 293-299. https://doi.org/10.1016/j.biombioe.2013.07.010

Padilla-Rivera, A., Paredes, M. G., & Guereca, L. P. (2019). A Systematic Review of the Sustainability Assessment of Bioenergy: The case of Gaseous Biofuels. *Biomass and Bioenergy*, 125, 79-94.

Martin-Gamboa, M., Iribarren, D., Garcia-Gusano, D., & Dufour, J. (2017). A Review of Life-Cycle Approaches coupled with Data Envelopment Analysis within Multi-criteria Decision Analysis for Sustainability Assessment of Energy Systems. *Journal of Cleaner Production*, 150, 164-174. https://doi.org/10.1016/j.jclepro.2017.03.017

Gabbay de Souza, R., Climaco, J. C. N., Sant'Anna, A. P., Rocha, T. B., Bastos do Valle, R., & Quelhas, O. L. G. (2016). Sustainability Assessment and Prioritisation of E-Waste Management Options in Brazil. *Waste Management*, 57, 46-56. https://doi.org/10.1016/j.wasman.2016.01.034

Akhoundi, A., and Nazif, S. (2018). Sustainability Assessment of Wastewater Reuse Alternatives using the Evidential Reasoning Approach. *Journal of Cleaner Production*, 195, 1350-1376.

An, D., Xi, B., Ren, J., Wang, Y., Jia, X., He, C., & Li, Z. (2017). Sustainability Assessment of Groundwater Remediation Technologies based on multi-Criteria Decision Making method. *Resources, Conservation and Recycling*, 119, 36-46. https://doi.org/10.1016/j.resconrec.2016.08.002

Irene De Luca, A., Iofrida, N., Leskinen, P., Stillitano, T., Falcone, G., Strano, A., & Gulisano, G. (2017). Life cycle Tools combined with Multi-Criteria and Participatory Methods for Agricultural Sustainability: Insights from a Systematic and Critical Review. *Science of The Total Environment*, 595, 352-370. https://doi.org/10.1016/j.scitotenv.2017.03.284

Joglekar, S. N., Kharkar, R. A., Mandavgane, S. A., & Kulkarni, B. D. (2018). Sustainability Assessment of Brick Work for Low-Cost Housing: A Comparison between Waste Based Bricks and burnt Clay Bricks. *Sustainable Cities and Society*, 37, 396-406. https://doi.org/10.1016/j.scs.2017.11.025

Yaylaci, E. D., and Duzgun, H. S. (2016). Indicator-based Sustainability Assessment for the Mining Sector Plans: Case of Afşin-Elbistan Coal Basin. *International Journal of Coal Geology*, 165, 190-200.

Munda, G. (2005). Multiple Criteria Decision Analysis and Sustainable Development. Multiple Criteria Decision Analysis: State of the Art Surveys, 78, 953-986. https://doi.org/10.1007/0-387-23081-5_23 "(quoted by Cinelli *et al*., 2014)".

EPA. (2006). Life Cycle Assessment: Principles and Practice. CSIRO Publishing. Cincinnati, OH: Office of Research and Development. "(quoted by Cinelli *et al*., 2014)".

Carayannis, E. G., Ferreira, J. J. M., Jalali, M. S., & Ferreira, F. A. F. (2018). MCDA in Knowledge-Based Economies: Methodological Developments and Real World Applications. *Technological Forecasting and Social Change*, 131, 1-3. https://doi.org/10.1016/j.techfore.2018.01.028

Belton, V., and Gear, T. (1983). On a Short-Coming of Saaty's Method of Analytic Hierarchies. *Omega*, 11(3), 228-230. https://doi.org/10.1016/0305-0483(83)90047-6

Perez, J., Jimeno, J. L., & Mokotoff, E. (2006). Another Potential Shortcoming of AHP. *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*, 14(1), 99-111.

Bana e Costa, C. A., and Vansnick. (2008). A critical analysis of the eigenvalue method used to derive priorities in AHP. *European Journal of Operational Research*, 187(3), 1422-1428.

Bystrzanowska, M., and Tobiszewski, M. (2018). How can Analysts use Multicriteria Decision Analysis? *Trends in Analytical Chemistry*, 105, 98-105. https://doi.org/10.1016/j.trac.2018.05.003

Velasquez, M., and Hester, P. T. (2013). An Analysis of Multi-Criteria Decision Making Methods. *International Journal of Operations Research*, 10(2), 56-66.

Ferretti, V., Bottero, M., & Mondini, G. (2014). Decision making and cultural heritage: An application of the Multi-Attribute Value Theory for the reuse of historical buildings. *Journal of Cultural Heritage*, 15(6), 644-655. https://doi.org/10.1016/j.culher.2013.12.007

Giuseppe, M. (1995). Multicriteria Evaluation in a Fuzzy Environment. Theory and Applications in Ecological Economics. Heidelberg, Germany: Physica-Verlag Springer.

Pohekar, S. D., and Ramachandran, M. (2004). Application of Multi-criteria Decision Making to Sustainable Energy Planning - A Review. *Renewable and Sustainable Energy Reviews*, 8, 365-381. doi:10.1016/j.rser.2003.12.007

Delforce, R. J., and Hardaker, J. B. (1985). AN EXPERIMENT IN MULTIATTRIBUTE UTILITY THEORY. *The Australian Journal of Agricultural Economics*, 29(3), 179-198. https://doi.org/10.1111/j.1467-8489.1985.tb00443.x

Ishizaka, A., and Siraj, S. (2018). Are Multi-Criteria Decision-Making Tools Useful? An Experimental Comparative Study of three Methods. *European Journal of Operational Research*, 264(2), 462-471. https://doi.org/10.1016/j.ejor.2017.05.041

Bana e Costa, C. A, and Chagas, M. P. (2004). A Career Choice Problem: An Example of How to use MACBETH to Build a Quantitative Value Model based on Qualitative Value Judgments. *European journal of Operational Research*, 153(2), 323-331. https://doi.org/10.1016/S0377-2217(03)00155-3

Dhouib, D. (2014). An Extension of MACBETH method for a Fuzzy Environment to Analyze Alternatives in Reverse Logistics for Automobile Tire Wastes. *Omega*, 42(1), 25-32.

Marques, R. C., Ferreira de Cruz, N., & Pires, J. (2015). Measuring the Sustainability of Urban Water Services. *Environmental Science & Policy*, 54, 142-151. https://doi.org/10.1016/j.envsci.2015.07.003

Lavoie, R., Deslandes, J., & Proulx, F. (2016). Assessing the Ecological Value of Wetlands using the MACBETH Approach in Quebec City. *Journal of Nature Conservation*, 30, 67-75.

Bana e Costa, C. A., Carnero, M. C., & Oliveira, M. D. (2012). A multi-criteria model for Auditing a Predictive Maintenance Programme. *European Journal of Operational Research*, 217, 381-393.

Bana e Costa, C. A. and Oliveira, R. C. (2002). Assigning priorities for Maintenance, Repair and Refurbishment in Managing a Municipal Housing Stock. *European Journal of Operational Research*, 138(2), 380-391. https://doi.org/10.1016/S0377-2217(01)00253-3

Bana e Costa, C. A., Lourenco, J. C., Chagas, M. P, & Bana e Costa, J. C. (2008). Development of Reusable Bid Evaluation Models for the Portuguese Electric Transmission Company. *Decision Analysis*, 5(1), 22-42. doi 10.1287/deca.1080.0104

Bana e Costa, C. A., De Corte, Jean-Marie., & Vansnick, Jean-Claude. (2003). MACBETH

Sanchez-Lopez, R., Bana e Costa, C. A., & De Baets, B. (2012). The MACBETH Approach for Multi-Criteria Evaluation of Development Projects on Cross-Cutting Issues. *Annals of Operations Research*, 199(1), 393-408. https://doi.org/10.1007/s10479-011-0877-4

IES. (2019). Website of Institute for Environmental Studies, http://www.ivm.vu.nl/en/projects/Archive/SustainabailityA-test/index.aspx, accessed on August 2nd.

Marttunen, M., Mustajoki, J., Dufva, M., & Karjalainen, T. P. (2013). How to Design and Realize Participation of Stakeholders in MCDA processes? A Framework for Selecting an Appropriate Approach. *EURO Journal on Decision Processes*, 3(1-2), 187-214. https://doi.org/10.1007/s40070-013-0016-3

Hage, M., Leroy, P., & Peterson, A. C. (2010). Stakeholder Participation in Environmental Knowledge Production. Futures, 42(3), 254-264. https://doi.org/10.1016/j.futures.2009.11.011

Ernst, A., Bib, K. H., Shamon, H., Schumann, D., & Heinrichs, H. U. (2018). Benefits and Challenges of Participatory Methods in Qualitative Energy Scenario Development. *Technological Forecasting and Social Change*, 127, 245-257. https://doi.org/10.1016/j.techfore.2017.09.026

Rowe, G. and Frewer, L. J. (2000). Public Participation Methods: A Framework for Evaluation. *Science, Technology, & Human Values*, 25(1), 3-29.

Elghali, L., Clift, R., Sinclair, P., Panoutsou, C., & Bauen, A. (2007). Developing a Sustainability Framework for the Assessment of Bioenergy Systems. *Energy Policy*, 35(12), 6075-6083.

Barfod, M. B., and Salling, K. B. (2015). A New Composite Decision Support Framework for Strategic and Sustainable Transport Appraisals. *Transportation Research Part A: Policy and Practice*, 72, 1-15.

Reddy, B. P., Kelly, M. P., Thokala, P., Walters, S. J., & Duenas, A. (2014). Prioritising public health guidance topics in the National Institute for Health and Care Excellence using the Analytic Hierarchy Process. Public Health, 128(10), 896-903. https://doi.org/10.1016/j.puhe.2014.07.003

Rodrigues, T. C. (2014). The MACBETH Approach to Health Value Measurement: Building a Population Health Index in Group Processes. *Procedia Technology*, 16, 1361-1366. https://doi.org/10.1016/j.protcy.2014.10.153

Van Schoubroeck, S., Springael, J., Van Dael, M., Malina, R., & Van Passel, S. (2019). Sustainability Indicators for Biobased Chemicals: A Delphi study using Multi-Criteria Decision Analysis. *Resources, Conservation and Recycling*, 144, 198-208. https://doi.org/10.1016/j.resconrec.2018.12.024

Doukas, H., Patlitzianas, K. D., & Psarras, J. (2006). Supporting Sustainable Electricity Technologies in Greece using MCDM. *Resources Policy*, 31(2), 129-136. https://doi.org/10.1016/j.resourpol.2006.09.003

Ivlev, I., Vacek, J., Kneppo, P. (2015). Multi-Criteria Decision Analysis for Supporting the Selection of Medical Devices Under Uncertainty. *European Journal of Operational Research*, 247(1), 216-228.

Vieira, A. C. L., Oliveira, M. D., & Bana e Costa, B. A. (2019). Enhancing knowledge construction processes within multicriteria decision analysis: The Collaborative Value Modelling framework.

Foth, T., Efstathiou, N., Vanderspank-Wright, B., Ufholz, L., Dutthorn, N., Zimansky, M., & Humphrey-Murto, S. (2016). The use of Delphi and Nominal Group Technique in Nursing Education: A Review. *International Journal of Nursing Studies*, 60, 112-120. https://doi.org/10.1016/j.ijnurstu.2016.04.015

Baron, R. S. (2005). So Right It's Wrong: Groupthink and the Ubiquitous Nature of Polarized Group Decision Making. *Advances in Experimental Social Psychology*, 37, 219-253.

Gardner, G., Jones, E., & Olesen, D. (1999). Use of Focus Group Technique to Investigate Infection Control Practice. *Australian Infection Control*, 4(4), 16-19. https://doi.org/10.1071/HI99416

Von der Gracht, H. K. (2012). Consensus Measurement in Delphi Studies Review and Implications for Future Quality Assurance. *Technological Forecasting & Social Change*, 79, 1525-1536.

Landeta, J. (2006). Current Validity of the Delphi Method in Social Sciences. *Technological Forecasting & Social Change*, 73 (5), 467-482. doi:10.1016/j.techfore.2005.09.002

Dalkey, N., and Helmer, O. (1963). An Experimental Application of the Delphi Method to the Use of Experts. *Management Science*, 9(3), 458-467. https://doi.org/10.1287/mnsc.9.3.458

Gnatzy, T., Wrath, J., Von der Gracht, H., & Darkow, I. (2011). Validating an innovative real-time Delphi approach - A methodological comparison between real-time and conventional Delphi studies. *Technological Forecasting & Social Change*, 78, 1681-1694. doi:10.1016/j.techfore.2011.04.006

Rowe, G., and Wright, G. (2001). Expert Opinions in Forecasting. Role of the Delphi Technique. In J. Armstrong (Ed.). Principles of Forecasting: A Handbook of Researchers and Practitioners. Boston, Massachusetts: Kluwer Academic Publishers.

Linstone, H. A., and Turoff, M. (1975). The Delphi Method: Techniques and Applications. Boston, MA: Addison-Wesley Educational Publishers Inc.

Reid, N. (1988). The Delphi Technique: Its Contribution to the Evaluation of Professional Practice. (Ellis R.) Professional Competence and Quality Assurance in the Caring Professions. London: Chapman and Hall.

Linstone, H. A., and Turoff, M. (2010). Delphi: A brief look backward and forward. *Technological Forecasting & Social Change*, 78, 1712-1719. doi:10.1016/j.techfore.2010.09.011

Coates, J. F. (1975). In Defense of Delphi: A Review of Delphi Assessment, Expert Opinion, Forecasting, and Group Process by H. Sackman. *Technological Forecasting and Social Change*, 7(2), 193-194. https://doi.org/10.1016/0040-1625(75)90058-X "(quoted by Linstone and Turoff, 2010)".

Belton, I., MacDonald, A., Wright, G., & Hamlin, I. (2019). Improving the Practical Application of the Delphi Method in Group-Based Judgment: A Six-Step Prescription for a Well-Founded and Defensible Process. *Technological Forecasting & Social Change*, 147, 72-82. https://doi.org/10.1016/j.techfore.2019.07.002

Phillips, L. D. (1984). A Theory of Requisite Decision Models. *Acta Psychologica*, 56(1-3), 29-48. https://doi.org/10.1016/0001-6918(84)90005-2

McCool, S. F., and Stankey, G. H. (2004). Indicators of Sustainability: Challenges and Opportunities at the Interface of Science and Policy. *Environmental Management*, 33, 294-305. doi: 10.1007/s00267-003-0084-4

Opon, J., and Henry, M. (2019). An Indicator Framework for Quantifying the Sustainability of Concrete Materials from the Perspectives of Global Sustainable Development. *Journal of Cleaner Production*, 218, 718-737. https://doi.org/10.1016/j.jclepro.2019.01.220

Goodman, C. M. (1987). The Delphi Technique: A Critique. *Journal of Advanced Nursing*, 12, 729-734.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research Guidelines for the Delphi Survey Technique. *Journal of Advanced Nursing*, 32(4), 1008-1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Watrobski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., & Ziolo, M. Generalised Framework for Multi-Criteria Method Selection. *Omega*, 86, 107-124. https://doi.org/10.1016/j.omega.2018.07.004

Bana e Costa, C. A., Ensslin, L., Cornea, E. C., Vansnick, Jean-Claude. (1999). Decision Support Systems in Action: Integrated Application in a Multicriteria Decision Aid Process. *European Journal of Operational Research*, 113(2), 315-335. https://doi.org/10.1016/S0377-2217(98)00219-7

Bana e Costa, C. A., and Beinat, E. (2005). Model-Structuring in Public Decision-Aiding.

Bana e Costa, C. A., and Vansnick, Jean-Claude. (1997). Applications of the MACBETH Approach in the Framework of an Additive Aggregation Model. *Journal of Multi-Criteria Decision Analysis*, 6, 107-114.

Bana e Costa, C. A., De Corte, Jean-Marie., & Vansnick, Jean-Claude. (2012). MACBETH. *International Journal of Information Technology & Decision Making,* 11(2), 359-387.

ENERPHI. (2019). Website of ENERPHI, https://enerphi.tecnico.ulisboa.pt/, accessed on August 28th.

Trevelyan, E. G., and Robinson, N. (2015). Delphi Methodology in Health Research: How to do it? European Journal of Integrative Medicine, 7(4), 423-428. https://doi.org/10.1016/j.eujim.2015.07.002

Fasolo, B, and Bana e Costa, C. A. (2014). Tailoring Value Elicitation to Decision Makers' Numeracy and Fluency: Expressing Value Judgments in Numbers or Words. *Omega*, 44, 83-90.

Belton, V., and Stewart, T. J. (2002). Multiple Criteria Decision Analysis: An Integrated Approach. Springer US: Kluwer Academic Publishers.

Appendix A

welphi

Dear Participant,

We would like to thank you for accepting our invitation to the web-based Delphi approach.

This questionnaire is being carried out on behalf of my master thesis in collaboration with PLANETIERS, a sustainable online marketplace (e-commerce platform), and IST (Instituto Superior Tecnico) to develop a multicriteria decision model to assess the sustainability of products accordingly to PLANETIERS' preferences.

The purpose of the Web-Delphi panel is to indicate their agreement or disagreement with the indicator's relevance in evaluating the sustainability of a product on the PLANETIERS online platform.

The responses and the comments posted will be kept completely anonymous. Each round is expected to take 20 min to complete.

Kindly follow these instructions for the 1st Round:

1. **Create Account and Login**
2. **Press "CONTINUE" to be directed to questionnaire**
3. **Indicate your agreement or disagreement to the following statement: "Is this indicator relevant in assessing the sustainability of a product?"**
4. **You can choose only one from a total of 6 choices; "*Strongly Disagree*", "*Disagree*", "*Neither Agree or Disagree*", "*Agree*", "*Strongly Agree*" and "*No Answer*".**
5. **Detailed description of each indicator can be viewed by clicking on the "Eye" icon.**
6. **You can choose the "No Answer" radio button in case the indicator does not fall under your field of expertise or do not wish to answer.**
7. You further have a possibility to justify your answers by commenting on each individual indicator by clicking the "Dialogue Box" icon.
8. Click "SAVE AND NEXT" to register your answers.

To login, you will be requested to create your own password for your email address [redacted]

To start your participation, you will be requested to visit the following website: app.welphi.com/versoes/v1/Pages/MessagePage.aspx?u=10440&p=4668&r=5788&d=5634&a=14242&d2=5635&eye=1&cmt=1

Please note the schedule for each round of the Delphi process:

1st Round: 18th of July – 23rd of July

2nd Round: 25th of July – 30th of July

3rd Round: 1st of August – 6th of August

Please feel free to contact if you have any questions through the following e-mail address: [redacted]

Best Regards,

Moeez Ahmed, M. Sc. Energy Engineering and Management at IST.

WISEDON • Lisbon • Portugal
For any questions contact us at support@welphi.com

Figure 25 - Delphi 1st Round Invitation Email

welphi

Dear Participant,

We would like to thank you for accepting our invitation to the web-based Delphi approach. Kindly find the instructions for the 2nd Round as following:

Kindly follow these instructions for 2nd Round:

1. Login
2. Press "CONTINUE" to be directed to questionnaire
3. Results obtained from the 1st Round will be presented alongside your responses in the 2nd Round
4. Shaded radio button showcases your answer from the previous round
5. The aggregate answers from all respondents is available in percentage or absolute votes for each indicator
6. You have an option to maintain or change your answers for each indicator.
7. You can change your response by pressing the "EDIT" button.
8. Click "SAVE AND NEXT" to register your answers.

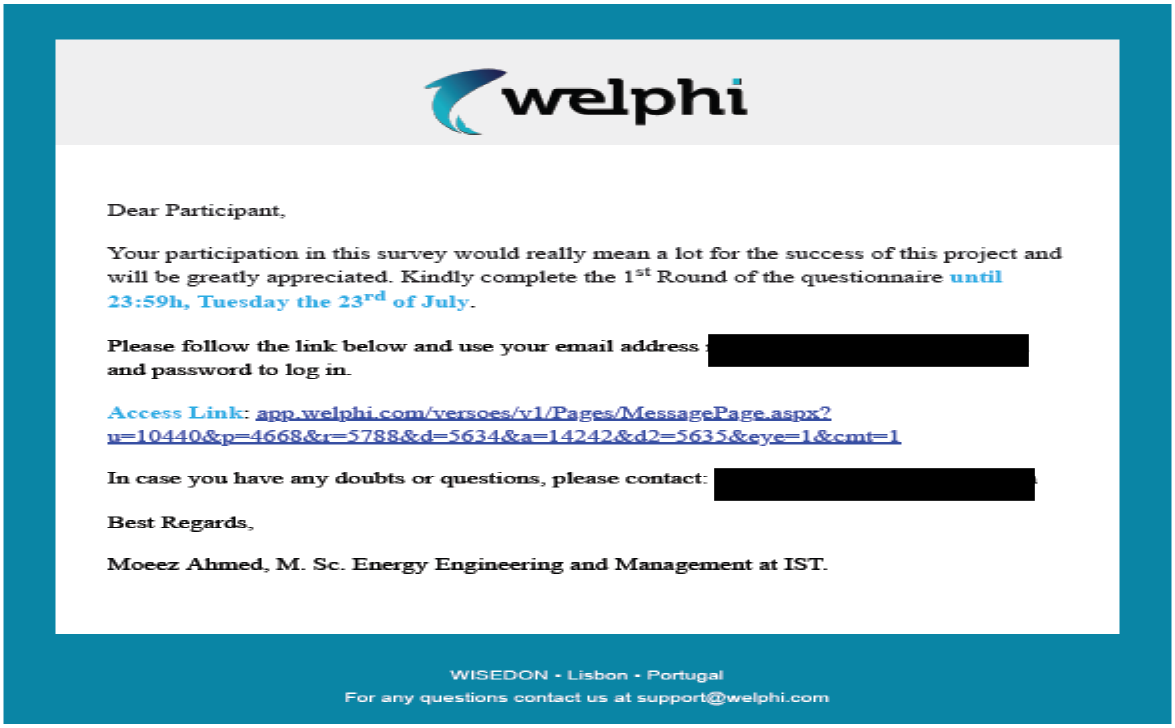
Please follow the link below and use your email address [redacted] and password to log in.

Access Link: app.welphi.com/versoes/v1/Pages/MessagePage.aspx?u=10440&p=4668&r=5788&d=5634&a=14242&d2=5635&eye=1&cmt=1

Please note the schedule for the 2nd Round of the Delphi process:

2nd Round: 26th of July – 31st of July

Figure 26 - Delphi Invitation Email 2nd round



welphi

Dear Participant,

Your participation in this survey would really mean a lot for the success of this project and will be greatly appreciated. Kindly complete the 1st Round of the questionnaire **until 23:59h, Tuesday the 23rd of July**.

Please follow the link below and use your email address [redacted] **and password to log in.**

Access Link: app.welphi.com/versoes/v1/Pages/MessagePage.aspx?u=10440&p=4668&r=5788&d=5634&a=14242&d2=5635&eye=1&cmt=1

In case you have any doubts or questions, please contact: [redacted]

Best Regards,

Moeez Ahmed, M. Sc. Energy Engineering and Management at IST.

WISEDON - Lisbon - Portugal
For any questions contact us at support@welphi.com

Figure 27 - Delphi Reminder Email

welphi

Dear Participant,

It would really mean a lot if you can spare 15 minutes to complete the survey. Your participation a lot for the success of this project and will be greatly appreciated. Kindly complete the 2nd Round of the questionnaire **until today 23:59h, 31st of July**.

Please follow the link below and use your email address [redacted] **and password to log in.**

Access Link: app.welphi.com/versoes/v1/Pages/MessagePage.aspx?u=10440&p=4668&r=5788&d=5634&a=14242&d2=5635&eye=1&cmt=1

In case you have any doubts or questions, please contact: [redacted]

Best Regards,

Moeez Ahmed, M. Sc. Energy Engineering and Management at IST.

WISEDON - Lisbon - Portugal
For any questions contact us at support@welphi.com

Figure 28- Delphi Last Reminder

Appendix B

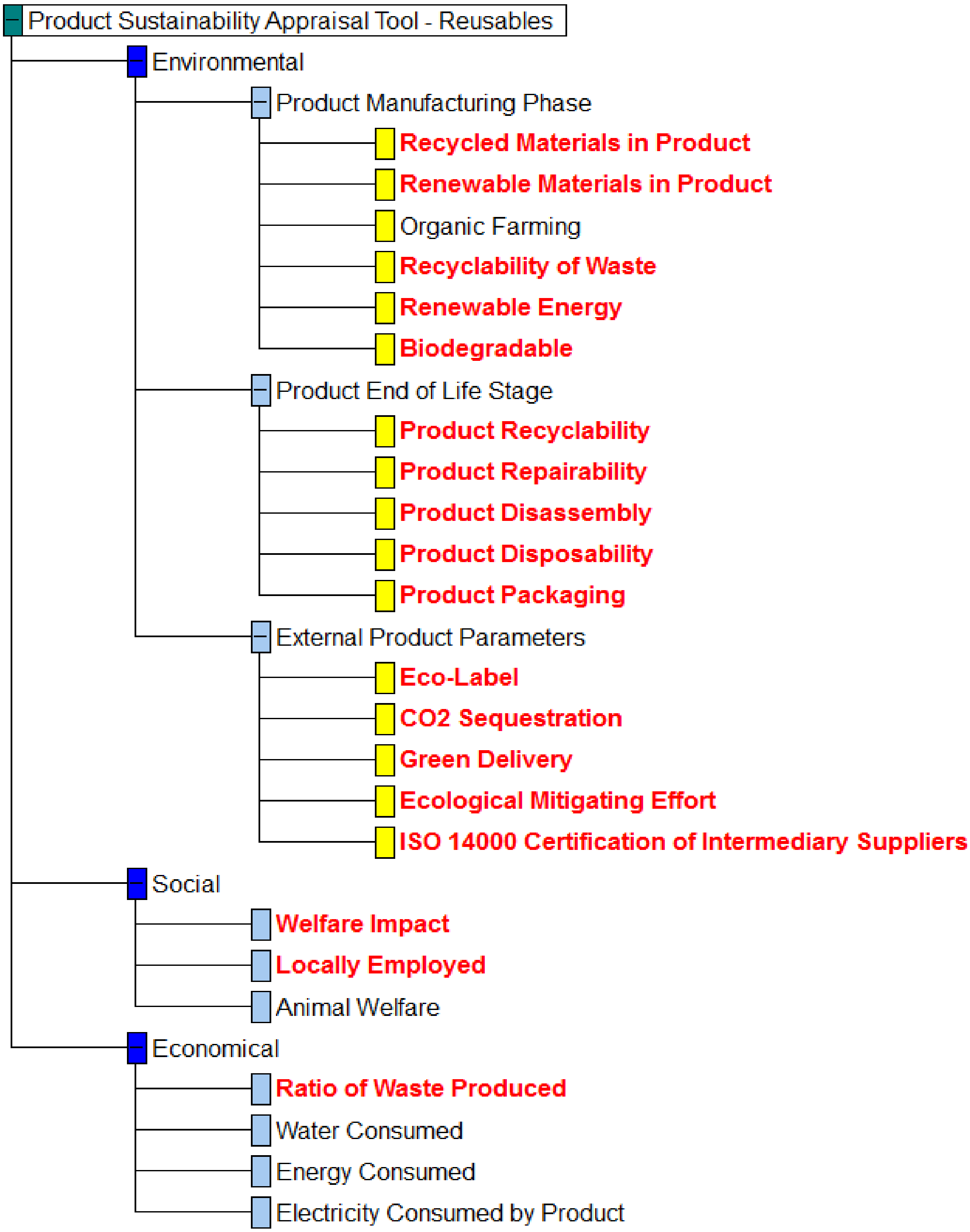


Figure 29 - Value Tree Reusables

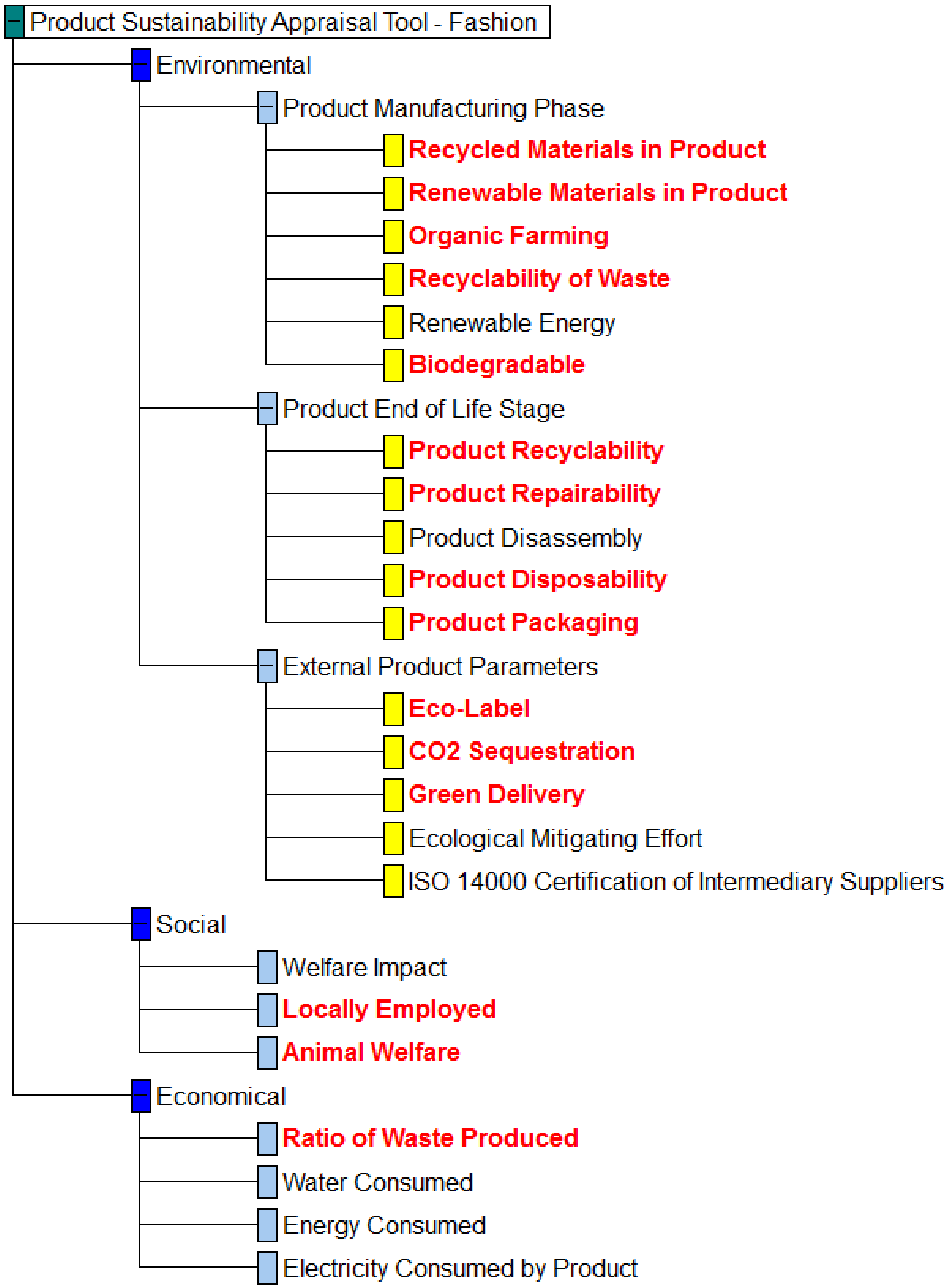


Figure 30 - Value Tree Fashion

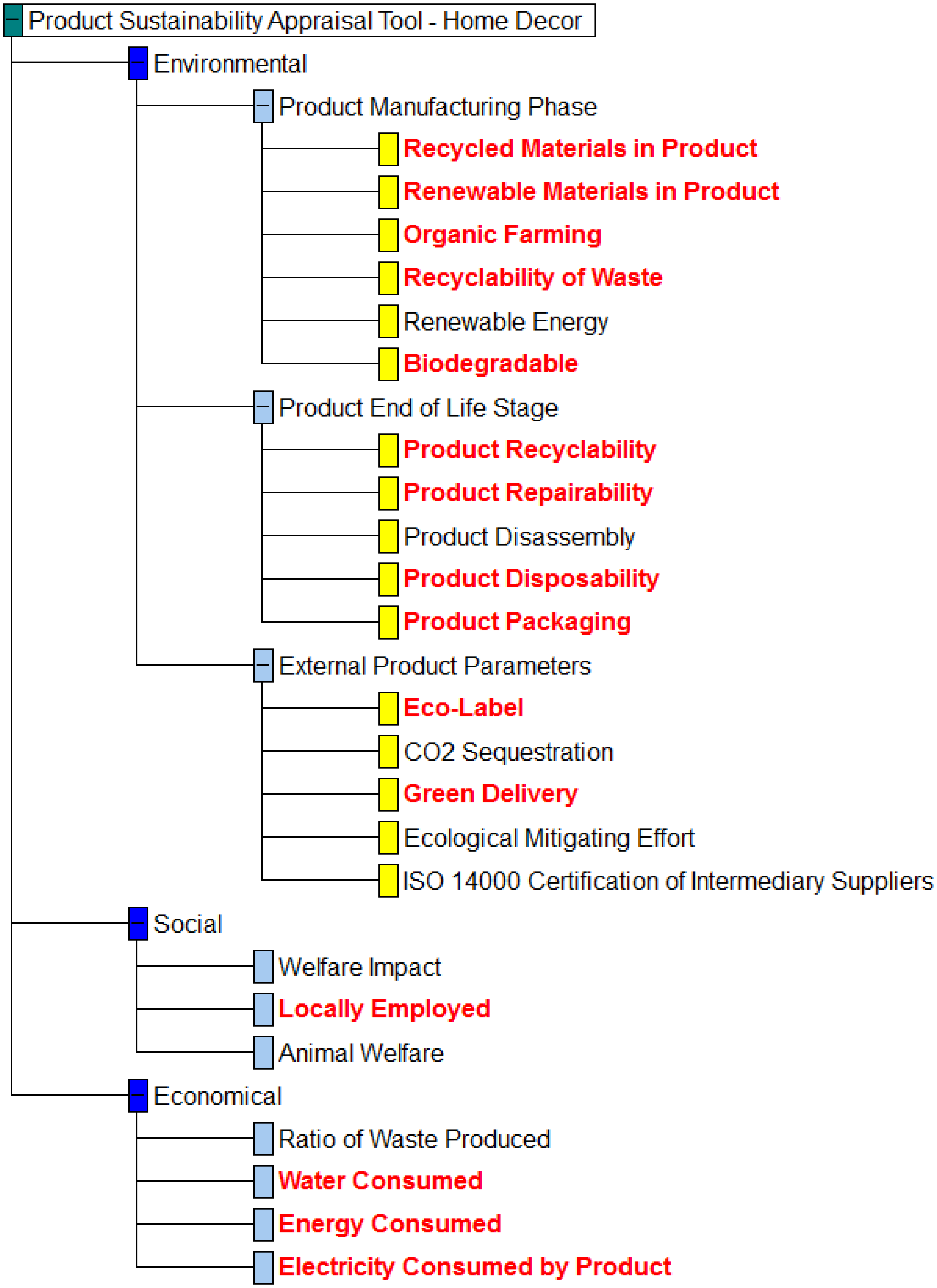


Figure 31 - Value Tree Home Décor

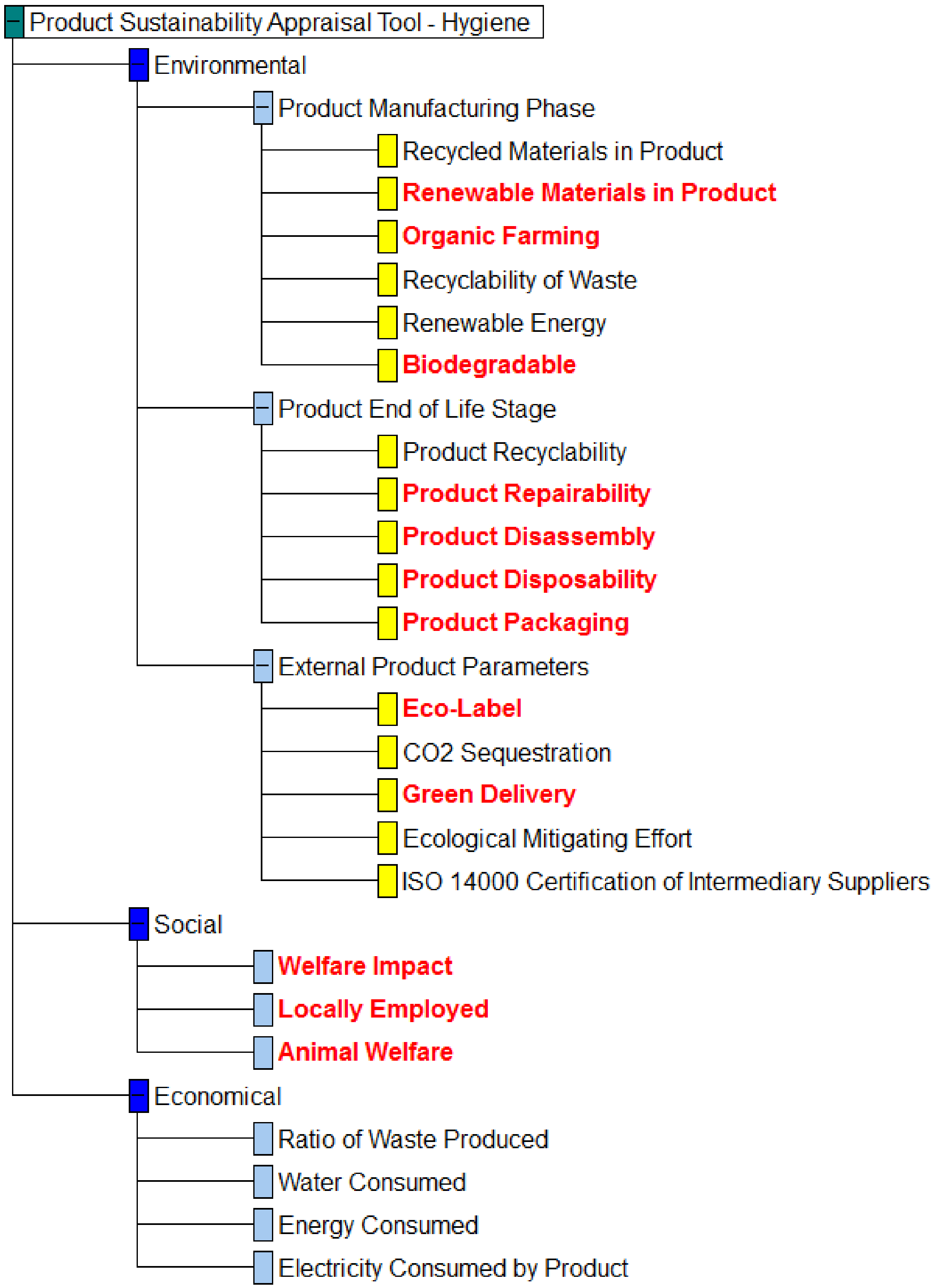


Figure 32 - Value Tree Hygiene

Appendix C

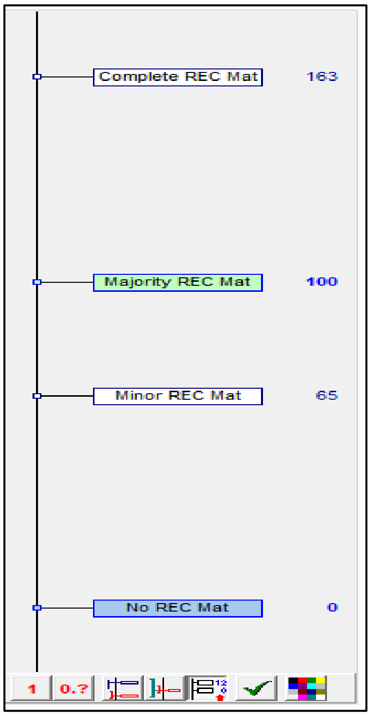


Figure 33 - Value Scale Recycled Materials in Product

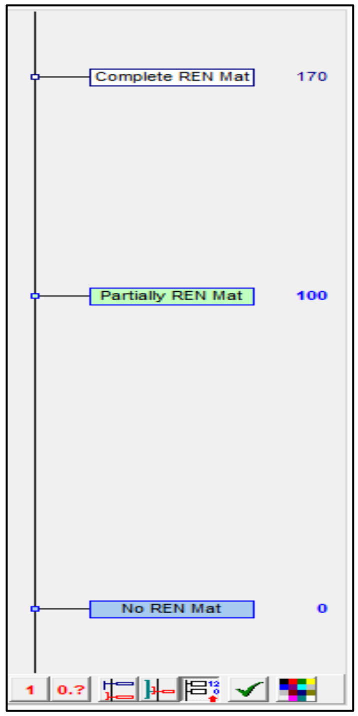


Figure 34 - Value Scale Renewable Materials in Product

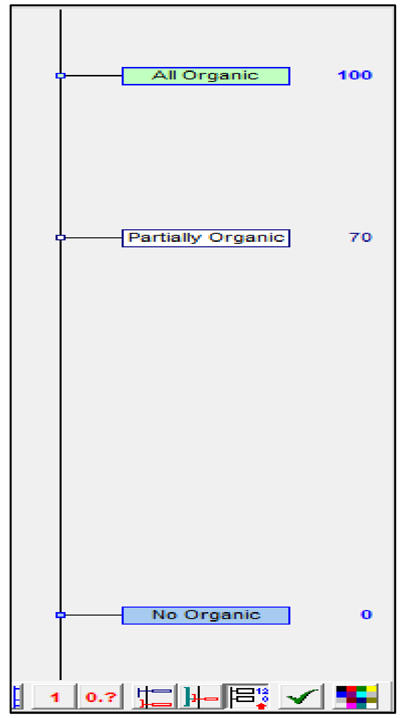


Figure 35 – Value Scale Organic Farming

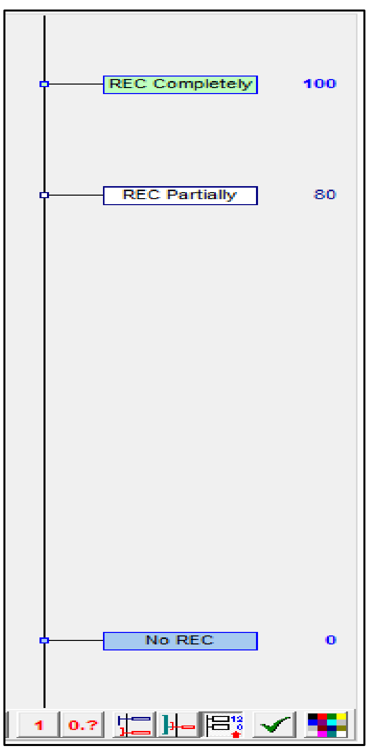


Figure 36 - Value Scale Recyclability of Waste

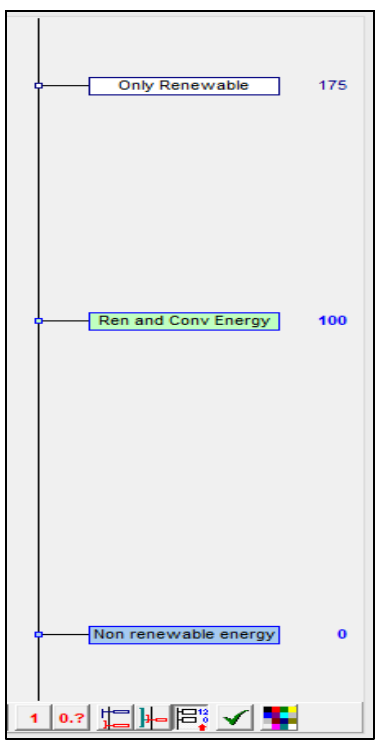


Figure 37 - Value Scale Renewable Energy

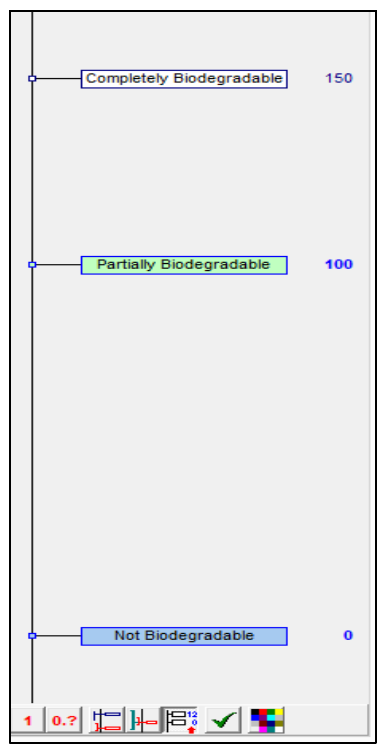


Figure 38 - Value Scale Biodegradable

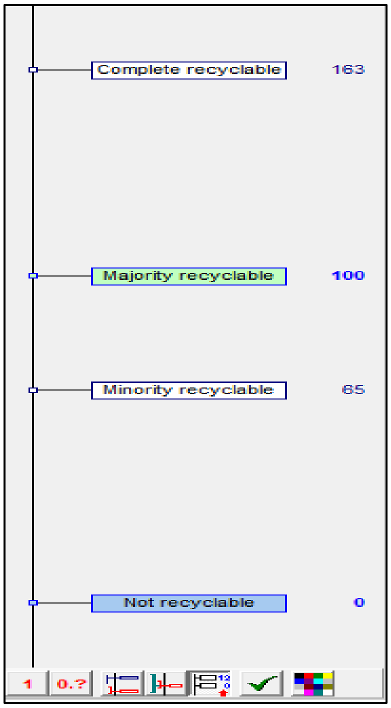


Figure 39 - Value Scale Product Recyclability

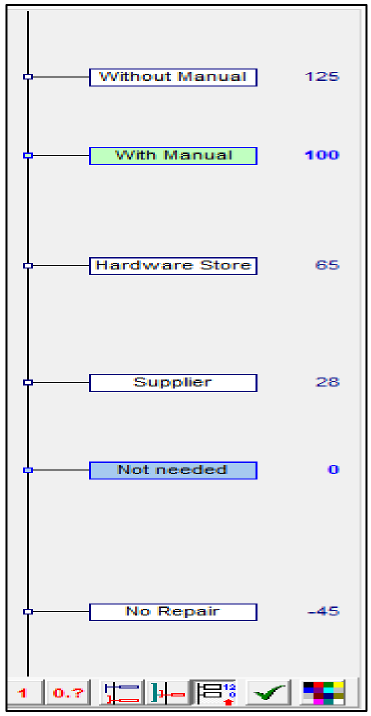


Figure 40 - Value Scale Product Repairability

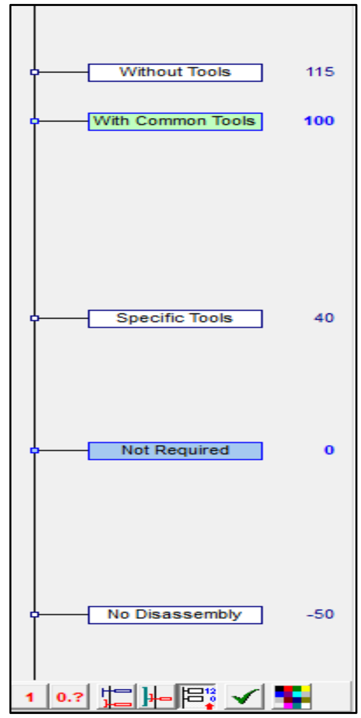


Figure 41 – Value Scale Product Disassembly

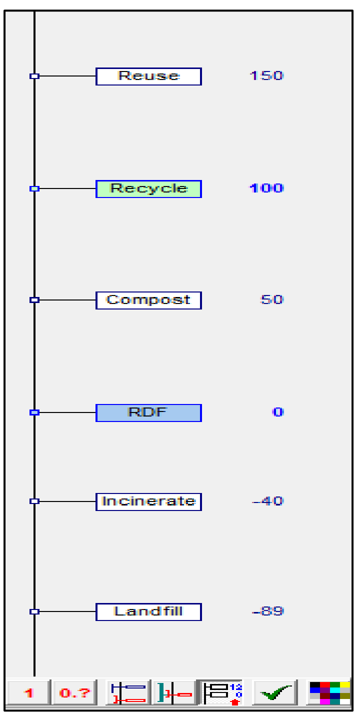


Figure 42 - Value Scale Product Disposability

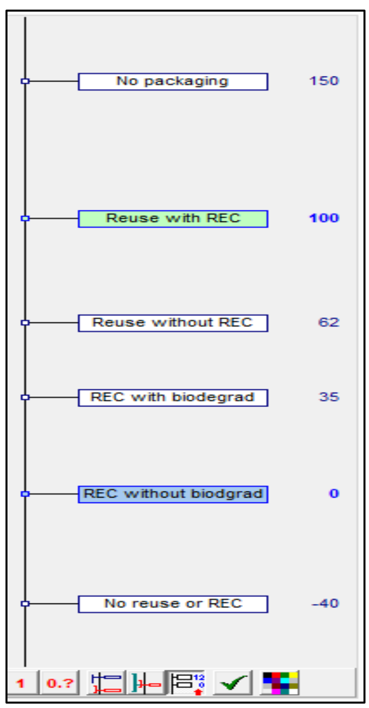


Figure 43 - Value Scale Product Packaging

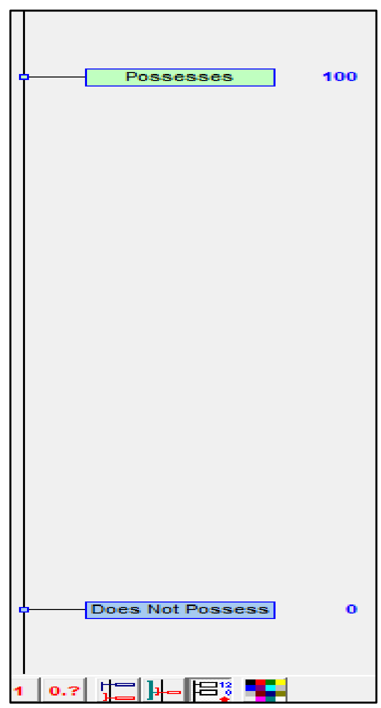


Figure 44 - Value Scale Eco-Label

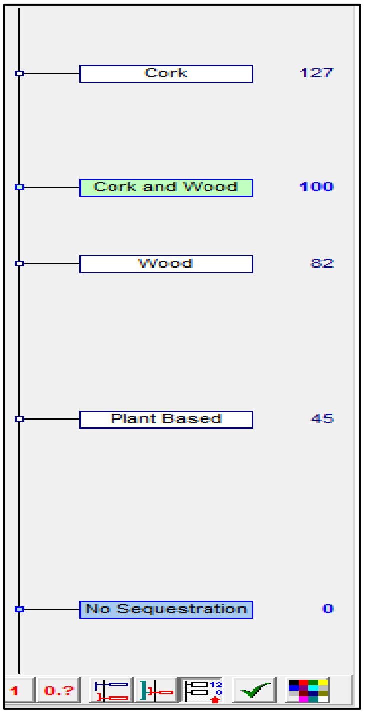


Figure 45 - Value Scale CO2 Sequestration

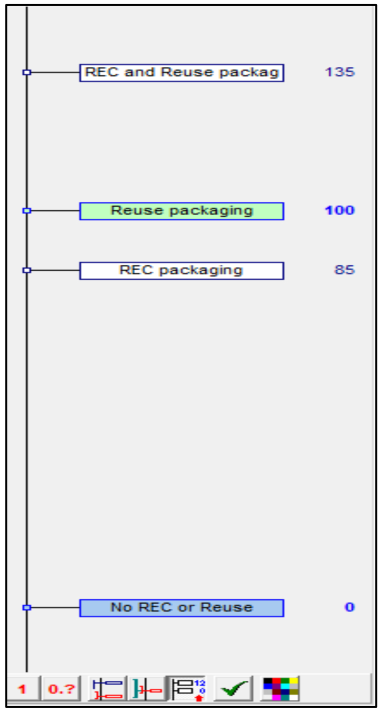


Figure 46 - Value Scale Green Delivery

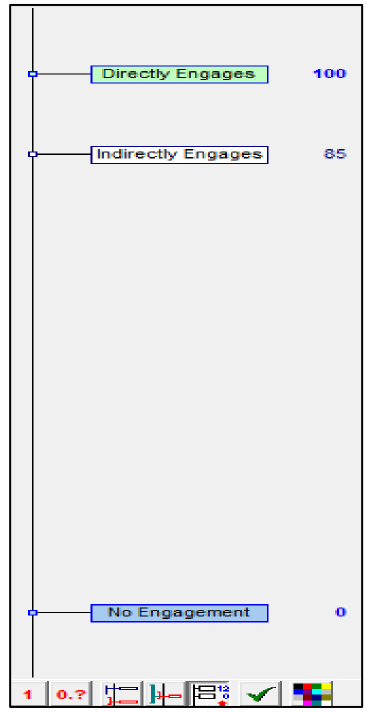


Figure 47 – Value Scale Ecological Mitigation Effort

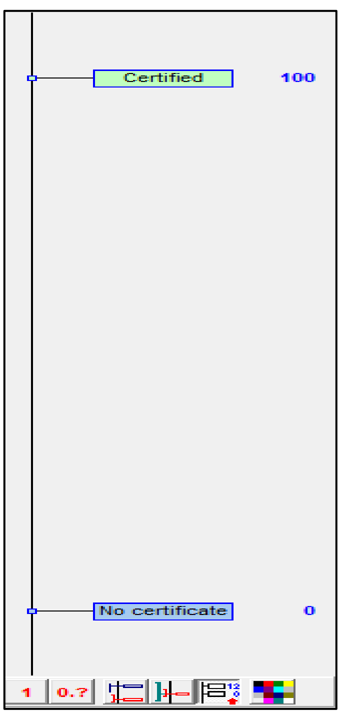


Figure 48 - Value Scale ISO 14000 Certification

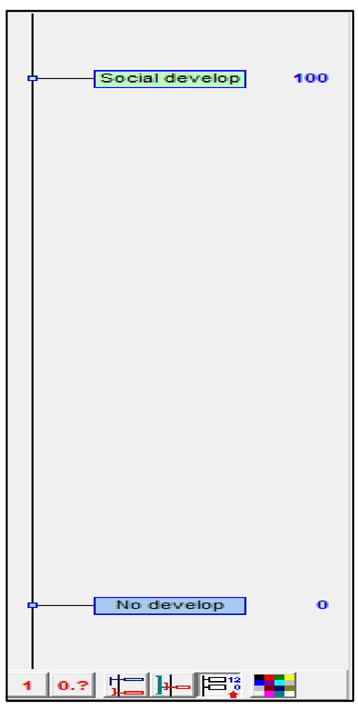


Figure 49 - Value Scale Welfare Impact

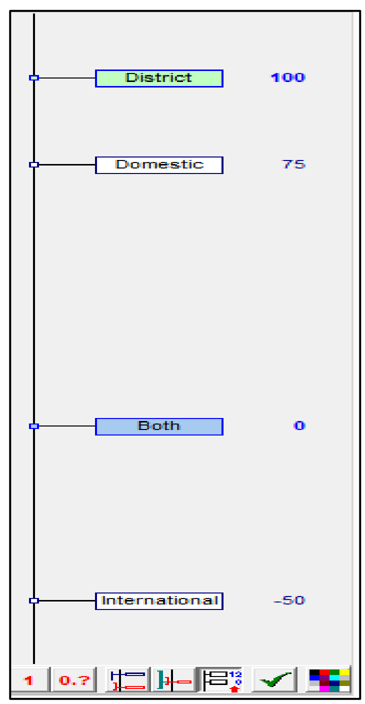


Figure 50 - Value Scale Locally Employed

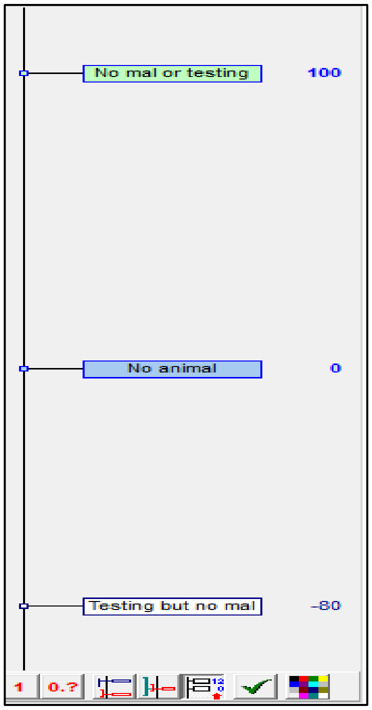


Figure 51 – Value Scale Animal Welfare

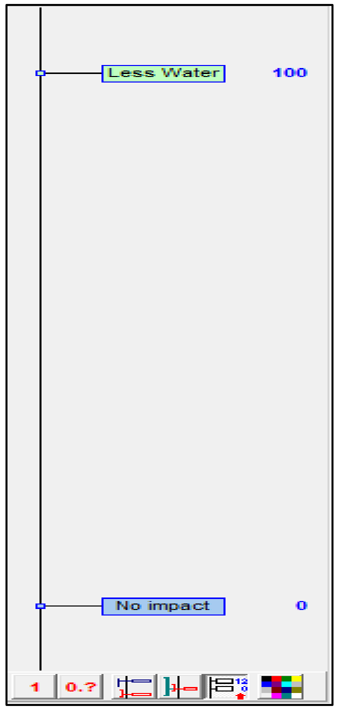


Figure 52 - Value Scale Water Consumed

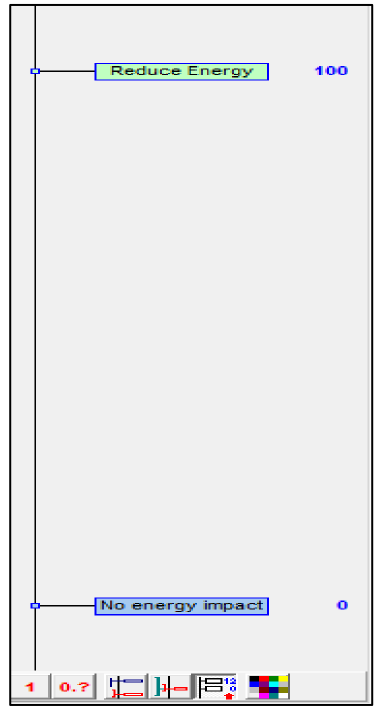


Figure 53 – Value Scale Energy Consumed

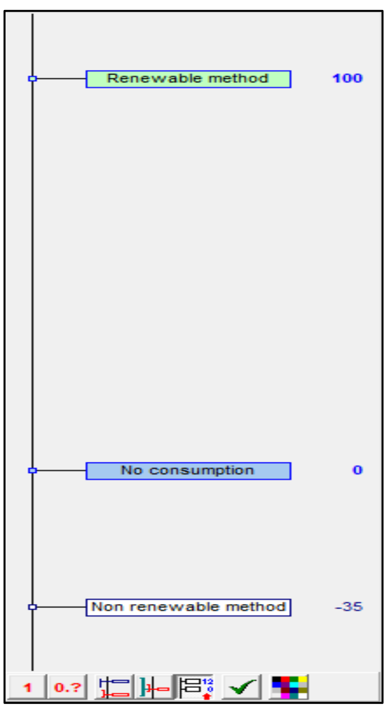


Figure 54 - Value Scale Electricity Consumed by Product

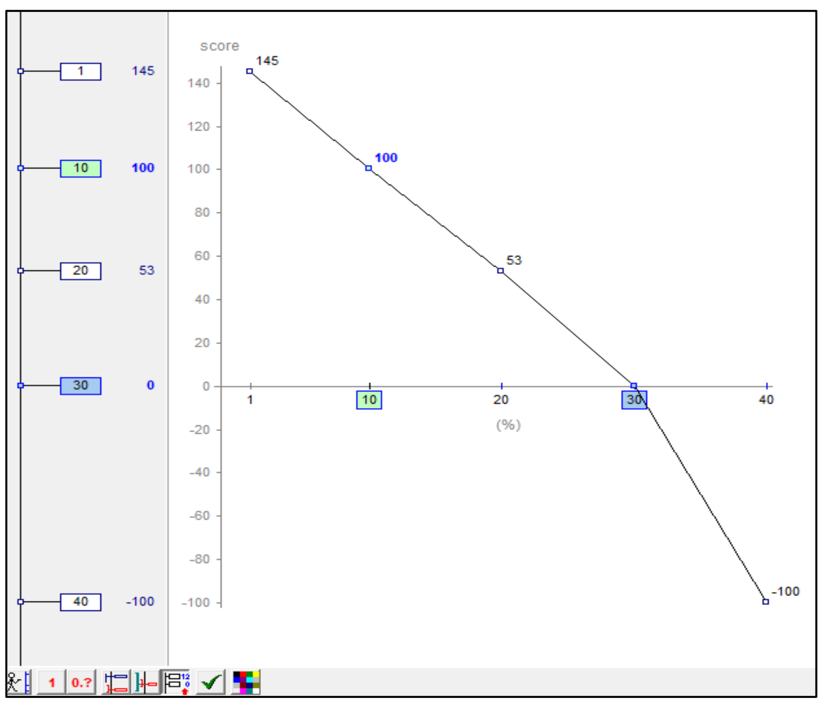


Figure 55 - Value Scale Ratio of Waste Produced

Appendix D

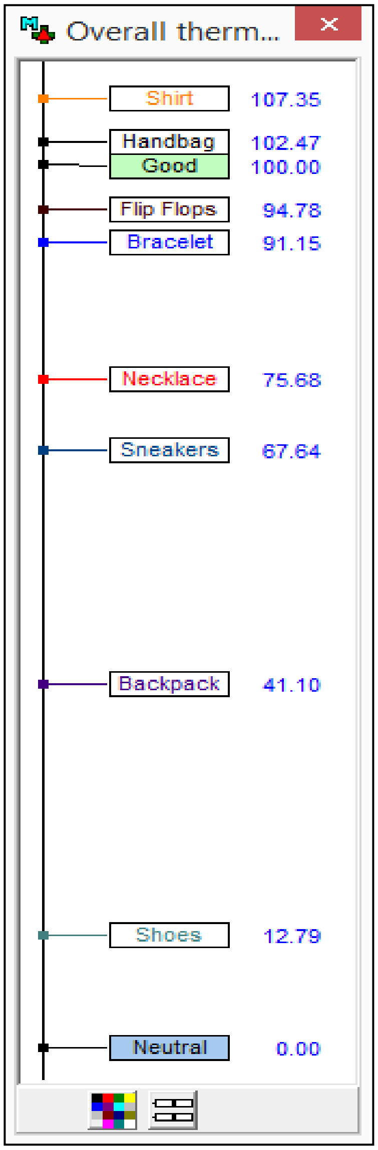


Figure 56 - Overall Thermometer Fashion

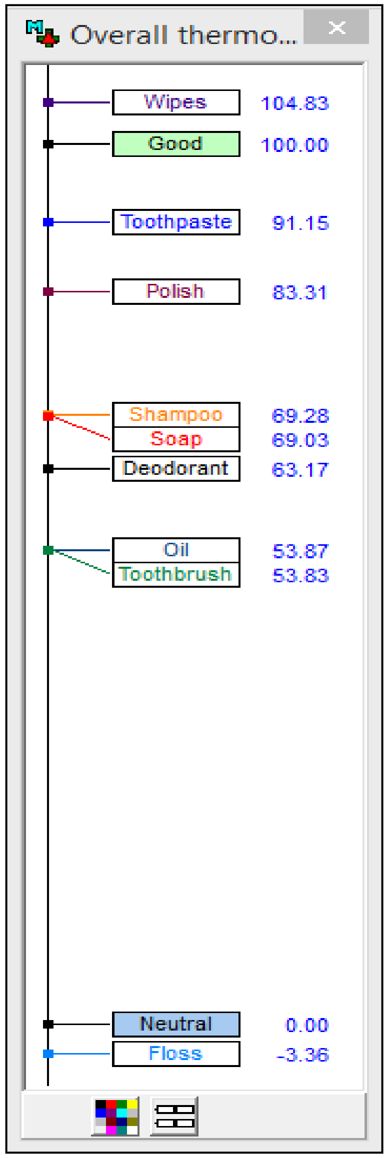


Figure 57 - Overall Thermometer Hygiene

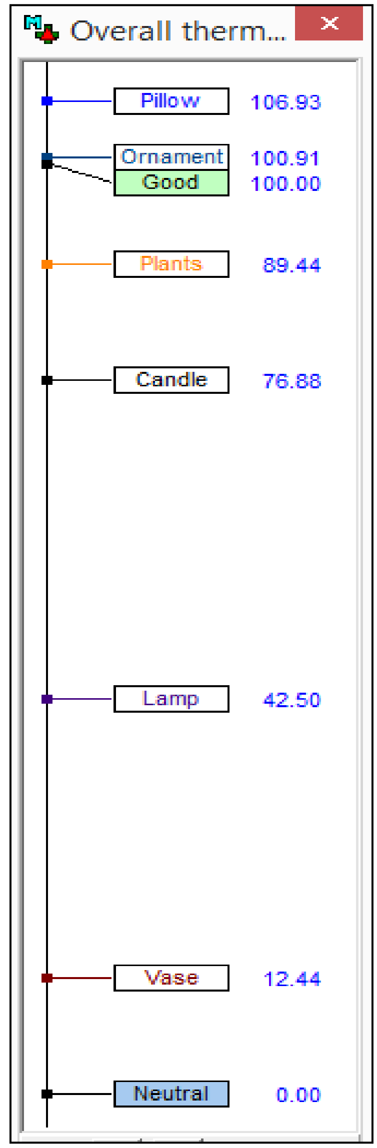


Figure 58 - Overall Thermometer Home Decor