The Need for a Preference-based Multicriteria Prioritization Framework in Life Cycle Sustainability Assessment

Emily Grubert

This is the pre-peer-reviewed version of the following article: Grubert, E. 2017. The Need for a Preference-based Multicriteria Prioritization Framework in Life Cycle Sustainability Assessment. Journal of Industrial Ecology, which has been published in final form at http://onlinelibrary.wiley.com/doi/10.1111/jiec.12631/full

Address correspondence to: Emily Grubert, Emmett Interdisciplinary Program in Environment and Resources, 473 Via Ortega, Y2E2 Building, Suite 226, Stanford CA, 94305 USA; Email: grubert@stanford.edu.

Keywords
life cycle assessment (LCA); environmental values; weighting; sensitivity analysis; decision support

Summary
Life cycle thinking is a valuable tool for integrated assessment of the environmental, social, and economic outcomes of human activities. The combination of the three as Life Cycle Sustainability Assessment (LCSA) is a powerful decision support tool, but it also presents important design challenges. Among the most important challenges is how to include subjective information necessary for defining the major elements of a decision: prospects to decide among, uncertainty, risk attitudes, and preferences. Previous work on values in life cycle methods has addressed prospects, uncertainty, and risk attitudes. This article builds on that work by arguing that given LCSA’s broad scope, explicit and standardized intercategory preferences are especially important for improving its value for decision makers. Practitioners should not be solely responsible for the value judgments necessary to integrate impact categories within and across environmental life cycle assessment (E-LCA), social LCA (S-LCA), and life cycle costing (LCC) evaluations for LCSA. Neither should this task fall entirely to decision makers, particularly as life cycle-grounded decisions are highly sensitive to value frames. Individuals are unlikely to be able to meaningfully interpret, evaluate, and determine tradeoffs without support. This article thus proposes that LCSA leverage its multiple paradigms to rigorously generate explicit, empirically grounded intercategory preference archetypes for use in evaluating decision robustness, much as cultural theory-based archetypes are currently used to test robustness to risk attitudes. Proof of concept data from the United States illustrate this approach, named WELFARES.

<heading level 1> Introduction
Life Cycle Assessment (LCA) is an important analytical tool for evaluating potential multicriteria outcomes associated with products, processes, and policies. LCA is traditionally applied to material flows from origin through use and disposal of all components of an output. Traditionally, these multicriteria outcomes have been negative and focused on the environment (Heijungs 1998), and LCA is often used to mean environmental LCA (E-LCA) (ISO 2006). LCA is one of the only tools designed with an aspiration toward holistic investigation of 1st through nth
order effects across the entire value chain of a desired output. Given interest in triple bottom line sustainability—which accounts for environmental, social, and economic outcomes—efforts are underway to develop an integrated life cycle tool that simultaneously addresses all three of these spheres (Guinée et al. 2011). This approach is sometimes referred to as Life Cycle Sustainability Assessment (LCSA), which in turn is often considered to integrate the three major existing life cycle methods: E-LCA, social LCA (S-LCA), and life cycle costing (LCC) (Jørgensen et al. 2013, Kloepffer 2008).

LCSA has methodological issues that require resolution, and major questions remain about whether LCSA is most appropriately considered a framework, a model, or something else. These challenges aside, many in the LCA and broader industrial ecology (IE) communities view LCSA as worth developing due to its potential usefulness as a decision support tool (Finkbeiner et al. 2010, Heijungs et al. 2010, Martínez-Blanco et al. 2014, Sala et al. 2013). The immediate implication is that the data and analytical frameworks that support LCSA must be accurate, replicable, fair, and complete. Completeness from a decision analysis perspective means addressing the major elements of a decision, including prospects (probabilistic alternative futures that involve risk, which map to the alternatives defined for producing a functional unit within LCA), uncertainties, risk attitudes, and preferences (see e.g. Kahneman and Tversky 1979, Hastie 2001, Howard 1988, Howard and Abbas 2015). Conducting clear, transparent, and rigorous analysis that explicitly engages relevant value judgments associated with these decision elements presents a major challenge (see e.g. Cowell et al. 2002, Finnveden 1999, Grubert 2016, Powell et al. 1997, Reap et al. 2008). Perhaps most importantly, there are not single, objectively correct approaches to making the many value judgments embedded in life cycle methods (see e.g. Finnveden 1997, Goedkoop et al. 1998, Hofstetter 1998, Hofstetter et al. 2000). Still, these value judgments have major implications for results (see e.g. De Schryver et al. 2009). Thus, this article argues that approaches to making value judgments should be explicit and standardized to increase transparency and comparability across life cycle studies.

One value judgment challenge that is particularly salient for LCSA given its broad scope is the question of how to decide between equally severe impacts in different impact categories—that is, the question of intercategory preferences. This article argues that LCSA needs an explicit and standardized way to integrate preferences across environmental, social, and economic impact categories. Explicitly accounting for intercategory preferences can strengthen LCSA as a decision support tool, joining existing approaches to defining decision prospects, uncertainty, and risk attitudes. In support of this argument, this article discusses the need for treatment of intercategory preferences within LCSA, outlines the use of empirically-grounded preference archetypes for sensitivity analysis (named WELFARES), and presents a proof of concept study using data from the United States.

<heading level 2> Defining Value Frameworks in a Multiparadigmatic Setting

LCSA is a potentially powerful decision support tool that is under active development, in large part due to the challenges of methodological, epistemological, and ontological blending. Relevant traditions include natural sciences and engineering, qualitative social sciences, and quantitative social sciences (e.g. economics) that traditionally address issues found in E-LCA, S-LCA, and LCC, respectively. Of the three individual life cycle methods, S-LCA is the least developed to date, with major outstanding questions about indicators, areas of protection, functional unit definitions, analytical targets, and others (Chhipi-Shrestha et al. 2015, Grubert 2016, Kloepffer 2008, Iofrida et al. 2016, Sala et al. 2013, Zanchi et al. 2016). LCSA inherits the
challenges of S-LCA (Martinez-Blanco et al. 2014) alongside further complexities associated with integrating it with E-LCA and LCC.

Perhaps the most fundamental challenge for S-LCA and its integration into LCSA is an issue common to interdisciplinary work: such integration merges traditionally positivist fields with traditionally interpretivist and constructivist fields (see e.g. Iofrida et al. 2016 for a deeper discussion and Allenby 2006 for more on the multiontological nature of industrial ecology). E-LCA and LCC are grounded in positivist paradigms that assert that one truth exists and can be at least partially described using experimental and observational methods. Though S-LCA is also grounded in historically quantitative and positivist fields (Sakellariou 2016), it addresses questions typically studied in interpretivist and constructivist fields that assert that multiple truths can coexist. Such questions include issues like whether an outcome is fair, whether a standard of living is adequate, or whether a change to the community is good. Conflict between S-LCA’s roots and goals contributes to the challenge of its development.

In general, making decisions requires integrating information about fundamentally noncomparable impact categories. Such integration requires value judgments, whether such judgments are made rigorously or not. An interpretivist frame is needed to incorporate the multiple value outlooks in any human society in order to make decisions informed by the reality that a given decision will affect people and systems differently. This acknowledgement of multiple subjective truths contrasts with the positivist frame historically used for indicator measurement, normalization, and characterization. Interpretivist paradigms lend themselves to consistent approaches to testing sensitivity to risk attitudes and intercategory preferences, where the existence of multiple valid perspectives is clear. Positivist paradigms lend themselves to inventory analysis and normalization procedures, where the quantitative LCSA method relies upon high quality measurements and consistency with observed causal pathways. Thus, LCSA can benefit from its blended paradigms.

LCSA provides an opportunity to broaden inquiry regarding careful ontological integration beyond questions raised by development of S-LCA. Each life cycle method includes both measurements and value judgments (including the value judgments inherent to measurement), and integrating expertise from different communities in appropriate areas can lend both power and systematization to the LCSA process (figure 1). The goal of creating an internally consistent and well-defined method for LCSA encourages the use of parallel processes across E-LCA, S-LCA, and LCC to the extent possible, which provides the benefit of improving each sub-method through LCSA development.

Questions about preferences within impact categories and elsewhere in life cycle methods are also germane, as are questions about normalizing data within and across impact categories. This article focuses on intercategory preferences largely because these pose challenges that are particularly relevant for LCSA relative to other life cycle methods. Given the theoretical grounding of E-LCA, S-LCA, and LCC in engineering and natural sciences, traditionally qualitative social sciences, and traditionally quantitative social sciences, respectively, LCSA faces questions not only of how weighting schemes are derived and applied but also of how to resolve interactions among vastly different epistemological and ontological frameworks. In essence, it requires addressing how information can be known and, more fundamentally, the nature of reality. Questions of how integrations and tradeoffs are derived and applied are already challenging, particularly because standard practice as defined for E-LCA by ISO 14044 proscribes the use of weighting schemes for comparative assertions presented to the public (2006). The gap between data and decisions imposes substantial cognitive burden by asking decision makers to make complex judgments without guidance, frequently resulting in a de facto equal weighting scheme or other arbitrary decision that can overwhelm even high quality LCA (for discussion of weighting schemes, see e.g. Finnveden et al. 2002, Johnsen and Løkke 2013, Pizzol et al. 2014, Weber and Borchering 1993). The need to enhance transparency about subjectivity and value choices and to better address decision maker needs is recognized in the literature (see e.g. Ehrenfeld and Gertler 1997, Finnveden 1999, Goedkoop et al. 1998, Hofstetter 1998, Johnsen and Løkke 2013, Nash and Stouthon 1994, Plevin et al. 2014a, Zhang 2010a), as is the need to better engage the public and include social issues (see e.g. Endter-Wada et al. 1998, Hofstetter et al. 2000, Johnsen and Løkke 2013, Thabrew et al. 2009).

**Decision Making in Life Cycle Methods**

Several distinct types of information support a decision, including prospects, uncertainty, risk attitude, and preferences. Prospects are relatively straightforward: these are the alternatives defined in the goal and scope stage of an LCA that a decision maker will choose among. Inventory analysis is used to characterize these prospects, including through normalization techniques that aim to put multicriteria impacts on equal footing (see e.g. Bare et al. 2006, Heijungs et al. 2007, Kim et al. 2013, Pizzol et al. 2016, Seppälä and Hämäläinen 2001, Sleeswijk et al. 2008). (Note, however, that intercategory normalization factors addressing environmental, social, and economic impact categories are not standardized.) During this characterization process, life cycle methods incorporate uncertainty related to measurement reliability, understanding of causal pathways, random fluctuations, and more. As with defining prospects, the way that life cycle methods incorporate uncertainty is relatively clear, despite implementation challenges. Less clear is the way risk attitude and preference information are supported within life cycle methods generally and LCSA specifically. Given the extreme sensitivity of LCA outcomes to risk attitude and preferences (see e.g. De Schryver et al. 2011, Grubert 2017a), careful consideration of these two decision elements is warranted.

Life cycle methods can be less effective for decision support when the responsible decision maker is unable to fully use the information included, a widely acknowledged challenge (see e.g. Rogers and Seager 2009, Seidel 2016, van Hoof et al. 2013). This is the impetus behind many efforts to build LCA weighting schemes. Even though LCSA is a decision support rather than a decision-making tool, its power is likely to be much greater if it includes guidance for decision makers concerning risk attitude and preferences. These elements are a major part of a
decision, and the complexity of a tool like LCSA means that incorporating these elements occurs at numerous stages of analysis.

The deep embedding of risk attitude in life cycle results is particularly salient, given its influence on issues like how to prioritize near- versus long-term outcomes, whether to prioritize more versus less certain outcomes, how much technological advancement to assume, and similar issues. One widely used approach to addressing risk attitude is through Cultural Theory (Hofstetter 1998, Hofstetter et al. 2000), which has been implemented in such impact assessment methods as Ecoindicator and ReCiPe (Goedkoop et al. 1998, Goedkoop et al. 2009). A key element of the Cultural Theory implementation is that multiple risk attitude archetypes are defined and consistently applied across studies, aided by software integration in much the same way that normalization references can be (see Heijungs et al. 2007 for a discussion of why software integration is important for assuring consistency).

Just as efforts have been made to standardize inventory data, normalization references, characterization factors (all addressing a decision prospect), inventory uncertainty, and risk attitudes through databases and software (see e.g. Bare 2002, Benoit Norris 2014, Frischknecht and Rebitzer 2005, Goedkoop et al. 1998, Huijbregts 1998, Jolliet et al. 2003, Lloyd and Ries 2007), so too do similar efforts to standardize preference data—especially regarding tradeoffs across impact categories in LCSA—stand to improve life cycle results’ value for decision making. To date, substantial work has been undertaken to address multicriteria decision analysis in the context of LCA (see e.g. Benoit and Rousseaux 2003, Gloria et al. 2007, Heijungs 1994, Hermann et al. 2007, Myllyviita et al. 2014, Seppälä 1999, Soares et al. 2006, Rogers and Seager 2009), but little work has explicitly sought to differentiate among risk, uncertainty, and preferences in LCA and LCSA. Addressing the questions of how to incorporate and how to standardize intercategory preference data in LCSA can potentially dramatically improve the tool’s value for decision makers by ensuring that all elements of a decision are included in LCSA.

**<heading level 3> The Decision Value of Preference Archetypes**

The primary contribution of this article is to propose that standardized intercategory preference archetypes be explicitly incorporated to LCSA, much as prospects, uncertainty, and risk attitudes already are. An extremely robust decision would appear favorable under all preference archetypes; a nonrobust decision would likely encourage more specific engagement with stakeholders on a project. Using empirically grounded archetypes rather than, for example, generating extreme conditions with a computer accounts for the possibility that there is general societal agreement on the relative importance of certain issues, even if specific priorities differ.

A major question about incorporating preference data to LCSA regards whose preferences should be incorporated. This work argues that preference archetypes should primarily be drawn from society at large, much as cultural theory draws risk attitudes from an assessment of the general population. Using archetypes derived from preferences in the general population does not mean fully democratizing LCSA. Rather, expert communities choose, characterize, normalize, and measure impact categories, and societal preferences are used to inform a course of action based on those results. Ultimately, decision makers decide.

The current approach to intercategory prioritization in life cycle-supported decision making is inadequate, as the nature of decision making means that such prioritization will occur with or without careful consideration or explicit statement of how. Results of a life cycle study are arguably more sensitive to the value choices used for multicriteria integration than to any other issue, and the value choices ultimately made overwhelm increased precision achievable.
through many of the methodological choices on which the community focuses more attention. Indeed, De Schryver et al. find that including Cultural Theory-based risk attitudes can alter characterization factors by six orders of magnitude (2011).

Even though implicit weighting is observable in the literature (Grubert 2017a), the community’s reluctance to define and apply explicit weighting schemes is understandable. LCA’s history as an industrial tool—and one explicitly used for policy making—means that concerns about liability remain salient (see e.g. Ashford 1997, Guinée et al. 2011). More generally, concern about the appropriateness of a small, unrepresentative group of primarily natural scientists and engineers making value judgments for a wider population is valid. Societally based preference archetypes thus provide additional value by enabling comparison with expert preferences like those elicited by Gloria et al. (2007) or implied by research focus areas (Grubert 2017a), highlighting areas where experts and society appear to disagree.

This article’s proposed framework for assessing a decision’s robustness based on empirically grounded data about societal preferences draws on the strengths of the multiple paradigms present in the LCA community. Further, by introducing preference archetypes as a set of numeric weighting factors that can be used with existing software tools, the path to a quantitative LCSA using standardized indicators, shared databases, and replicable methods is preserved. This work presents proof of concept data, but a more generalized and usable set of preference archetypes depends on agreement within the LCSA community on issues like which impact categories LCSA addresses and how preference measurements should be made, which requires ongoing multidisciplinary collaboration.

### Framework Description

In response to challenges outlined in the previous section, this article argues that LCSA needs an explicit and standardized way to integrate preferences across environmental, social, and economic impact categories if it is to be an effective decision support tool. The idea of a generally applicable set of societally-grounded intercategory preference archetypes for decision sensitivity analysis suggested by this article is henceforth referred to as WELFARES, an acronym loosely based on a set of environmental, social, and economic impact categories addressed in LCSA (figure 2). The success of WELFARES relies on both the preference archetypes themselves and on the strength and adaptability of elicitation methods. Shifting visions of which impact categories comprise LCSA, interest in generating both generic and case-specific preference archetypes, and a desire to catalog culturally specific preference data in a database (much as location-specific inventory data are cataloged) require that the archetypes not be static. The remainder of this article presents a proof of concept study conducted in the United States and discusses challenges and future efforts.

### Priority Archetypes for Life Cycle Sustainability Assessment

WELFARES builds on other work addressing the use of decision analysis, preference elicitation, and preference aggregation tools in LCA and related methods, for example Multicriteria Decision Analysis (MCDA) (e.g. Esteves 2008, Hertwich and Hammitt 2001, Linkov and Seager 2011); the Analytical Hierarchy Process (AHP) (e.g. Pineda-Henson et al. 2002, Saaty 2006); choice experiments for preference elicitation (e.g. Adamowicz et al. 1998, Mettler et al. 2006, Pizzol et al. 2014); and participatory decision making, including panels, surveys, and interviews (e.g. Gloria et al. 2007, Goedkoop et al. 1998, Mathe 2014, Soares et al. 2006).
Conceptually, WELFARES contributes to this literature by proposing three major characteristics for an intercategory preference weighting process for LCSA:

- An intercategory preference weighting scheme should be empirically grounded and should include preferences elicited from society at large;
- It should be based on preferences for environmental, social, and economic impact categories that are elicited together, such that a preference for any impact category included in LCSA can be directly compared with a preference for any other (which also enables use of the framework for E-LCA or S-LCA alone);
- It should include multiple preference archetypes, with a standard set of archetypes prioritizing preference diversity recommended for use with all LCSA studies (likely through software integration), and with flexibility to include an additional case-specific archetype generated from the relevant stakeholder groups using the same elicitation methods.

WELFARES is a nonmonetary weighting framework based on societal rather than expert priorities. It is intended for use as an impact category-level priorities-based sensitivity analysis tool (adding to the contributions of expert-based weighting schemes by e.g. Eco-indicator 99, Gloria et al. 2007, and Soares et al. 2006). The weighting schemes intentionally reflect an interpretivist view of the world as containing multiple truths about what is most important, intended to inform an answer, as opposed to proposing a single “most correct” weighting intended to produce an answer. This goal of analyzing sensitivity to intercategory preferences adds to work investigating other aspects of valuation, including sensitivity to risk attitude (Bare and Gloria 2006, De Schryver et al. 2009, Eco-indicator 99, Hellweg et al. 2003, Tukker 2002), impact monetization (Ahlroth 2014, Ahlroth et al. 2011, Finnveden et al. 2009, Hellweg et al. 2003, Jeswani et al. 2010, Pizzol et al. 2014, Weidema et al. 2013, Weidema 2016), and others. WELFARES is particularly intended for the LCSA context because of its explicit focus on generating an internally consistent weighting scheme designed to address environmental, social, and economic impacts.

<heading level 3> A United States-based WELFARES Proof of Concept

As a WELFARES proof of concept, priority rankings for 14 impact categories encompassing environmental, social, and economic issues (table 1) were elicited using a single-level AHP-style pairwise ranking questionnaire. The simplified list of 14 impact categories was adapted from larger lists of categories used in the E-LCA and S-LCA literatures, with the intent of capturing high-level issues. The use of a simplified list was intentional, given that impact categories in LCSA need to be further standardized before a fully usable set of WELFARES archetypes can be generated. Language was piloted before the study was launched to ensure that respondents generally interpreted the descriptions of an impact similarly. A pairwise ranking approach was chosen because it enables cardinal ranking and direct analysis of any subset of included characteristics.

Participants were adult United States residents recruited through two microtask platforms (Amazon Mechanical Turk and Prolific Academic) to complete a Qualtrics-based web questionnaire, with the study designed in accordance with Dynamo’s “Guidelines for Academic Researchers” (Dynamo 2015). At the beginning of the questionnaire, participants were randomly sorted into three prompt groups designed to incorporate intergenerational equity considerations to the data by asking participants to consider what they would prioritize in a community where their parent, they, or their child is moving. Questionnaire items displaying two topics and asking...
participants “Which is higher priority?” and “How much higher priority?” were presented one at a time (see supporting information, SI). The web questionnaire format allowed for randomization of the order of both items and impact categories within an item (that is, some participants saw “A) Human health or B) Ecosystem health,” whereas others saw “A) Ecosystem health or B) Human health”). In total, 21,000 pairwise comparisons were completed by 500 participants (17,000 by 200 Mechanical Turk Workers and 4,000 by 300 Prolific Academic participants) whose locations roughly correspond to a population map of the United States (figure 3). Participants recruited from Mechanical Turk and Prolific Academic tend to be better educated, younger, more politically liberal, and less racially diverse than the general population (Ipeirotis 2010, Paolacci and Chandler 2014, Prolific Academic 2016).

Results were coded and analyzed using the AHP method for a single level study and aggregated based on Aggregation of Individual Priorities, as participants are not assumed to behave as a single unit (Ossadnik et al. 2016). Results were also tested using Aggregation of Individual Judgments, which produced very similar cardinal rankings. Clustering via \(k\)-means \((k = [2,10])\) was performed on the combined Mechanical Turk and Prolific Academic sample to prioritize archetype diversity. Results for \(k = 5\) are presented in figure 4 and tables 2 and 3, as five clusters reflect the twin goals of effectively capturing pattern diversity and keeping the number of archetypes low enough to be useful for sensitivity analysis. This judgment is based on the presence of a quantitatively observable elbow in both within- and between-class variance between \(k = 4\) and \(k = 6\) (figure S1) and a subjective evaluation that distinct and human-parseable patterns that persist through higher values of \(k\), like an archetype prioritizing water or one prioritizing lifestyle, are clearly visible in the \(k = 5\) results. Results for other values of \(k\) are available in the SI (tables S1-S9).

Overall, proof of concept study results indicate relatively good agreement about which issues are more versus less important but still reveal distinct preferences (see table 3 for Spearman’s rho values relative to the aggregate ranking). A subjective evaluation suggests a focus on safety, water, health, personal quality of life, and community for Archetypes 1-5, respectively (figure 4, tables 2-3). An additional observation is that the most distinctly community oriented archetype (5) approaches a relatively equal weighting across all impact categories.

Ordinal results are more similar than cardinal results. With two exceptions, the top five priorities are the same for all five clusters and the aggregate estimate (table 3), though the associated cardinal values vary substantially (table 2). Similarly, the bottom five priorities are similar across the archetypes, with the notable exception of Archetype 5’s emphasis on community future (table 3).

Results also suggest that the frame of an intergenerational anchor affects participants’ responses, but not so much that underlying preference patterns are overwhelmed. That is, every preference archetype is present to some extent for every intergenerational anchor. The largest deviation from random distribution is that participants anchored on preferences for a parent versus for themselves heavily favored Archetype 1 over Archetype 4, and vice versa (figure 5). This pattern makes sense, as Archetype 1 prioritizes safety, whereas Archetype 4 prioritizes wages. These archetypes are derived from a relatively small, US-based sample and are thus unlikely to be either nationally or globally representative, but the emergence of well distributed, reasonable patterns from the sample data is encouraging.

<heading level 3> Methodological Development for WELFARES
Generating and updating data is a persistent challenge for life cycle methodologies, and expectations for continued changes in LCSA—particularly related to choice of impact categories—suggest that this challenge will be particularly relevant for LCSA. Standardizing a preference framework like WELFARES relies on the assurance that appropriate preference data will be available as LCSA matures and elicitation strategies improve. That is, a static WELFARES addressing the wrong impact categories or based on outdated elicitation techniques is not useful. Just as inventory data are centralized and updated as conditions, measurement capabilities, and LCA approaches change, so too should preference data be able to adapt to changing realities for LCSA. Thus, particularly given that similar efforts have frequently failed to be widely deployed, WELFARES must be a mutable framework based on a standardized process and centralized access (e.g. through software).

As with many other data-related tools, cost and maintenance pose large challenges for WELFARES. Data collection for the proof of concept study described in this work cost about $0.05 per pairwise comparison in payments to participants alone. Results from a broader mail-based survey (with a goal of reaching populations difficult to recruit online) assessing 30 impact categories suggest costs of between $20 and $40 per response for much lower resolution data, relying on Likert-type items rather than pairwise comparisons (Grubert 2017b). Anticipated costs for a broader study addressing a full complement of about 30 impact categories (necessitating 435 pairwise comparisons without fatigue, a major cognitive challenge), using a sample size of about 100,000 people globally to capture rare perspectives and provide sufficient resolution to assess cultural trends, would likely be millions of dollars, not including costs associated with recruiting representative samples, convening data quality verification meetings, or analyzing the data. Gathering data on preferences can be meaningfully costly, especially if a lack of methodological stability in LCSA means that data collection needs to be repeated over time. Thus, developing effective and preferably low-cost tools for maintaining WELFARES archetypes is a priority.

To date, two major experiments focused on data collection and maintenance within WELFARES have been completed or are underway. The first, associated with this proof of concept study, demonstrates that risks associated with cognitive overload (see e.g. Mettier et al. 2006) in a many-pair pairwise comparison study can be mitigated by asking each respondent to answer a selection rather than the full complement of pairwise comparison questions. This finding is valuable in contexts like this where a community or societal perspective is sought rather than an individual’s perspective. Mechanical Turk participants responded to 91 pairwise comparisons, whereas Prolific Academic participants each responded to a randomly selected set of 13: overall results for the two datasets were similar ($R^2 = 0.95$). Data quality for the aggregated dataset is acceptable, though it is difficult to derive archetypes from a split dataset. For practitioners interested in complementing a standardized set of WELFARES preference archetypes with case-specific archetypes, this split data generation approach could be a useful way to derive case-specific archetypes relatively quickly. A second experiment (Grubert in prep.) focuses on the use of digital tools to replicate data derived from a survey instrument. This ongoing work attempts to validate low-cost computational methods of deriving preference data from existing written documents, which would also allow for retrospective analysis, highly temporally and spatially specific analysis, and more robust archetypes. Such a digital method would improve LCSA practitioners’ ability to update WELFARES archetypes based on the tool’s development stage.
Discussion

Using empirically grounded intercategory preference archetypes to test decision robustness is a useful opportunity to consider what LCSA aims to accomplish and where different branches of science and inquiry can best contribute. Fundamentally, decisions about tradeoffs among various social and environmental indicators rely on the application of subjective value judgments at multiple junctures, including the basic paradigm of the work, choice of impact categories, derivation of characterization factors, and many more (see e.g. Allenby 2006, Ayres 1995, Hertwich and Hammitt 2001, Hauschild 2005, Souza et al. 2014, Zhang 2010b). Who makes those value judgments, and whose value systems are considered during the decision process, is a matter of some debate (see e.g. Endter-Wada et al. 1998, Johnsen and Løkke 2013, Thabrew et al. 2009, van den Bergh and Verbruggen 1999). Experts can benefit from understanding how acting on value systems they do not personally hold can affect decisions. This is particularly relevant given that ecological science has a history of viewing people as obstacles to decisions and as sources of damage to rather than beneficiaries of natural resources (Endter-Wada et al. 1998) and given a real chance that the expert community holds nonrepresentative ethical orientations that make other perspectives difficult to anticipate (Hauschild 2005).

A fully implemented WELFARES provides value by defining a process for eliciting empirically grounded intercategory preferences, then providing archetypical priority patterns that LCSA practitioners can use to test decision robustness and likely points of conflict. A major implementation concern that applies to any elicitation-based process regards issues of sampling and instrument design to ensure reliable measurements through, for example, appropriate framing, adaptation to cultural contexts, unambiguous question design, and others (e.g. Kahneman and Knetsch 1992, Mettier et al. 2006, Mettier and Scholz 2008, Mettier and Hofstetter 2004). An immediate concern more specific to WELFARES is whether applying societally derived preferences either elevates or effaces community input to an unacceptable degree. Does such a framework give society too much power? One might argue, for example, that incorporating community-based priority data could lead to immoral or scientifically unsupported recommendations (see e.g. Maier 2012 for a critique of preference-based decision making, especially Chapter 2.1.2.1, and Johnsen and Løkke 2013 for a discussion of metrics of “good” weighting systems prioritized by scientists versus society).

Three major responses stand out. First, WELFARES is intended to reflect a range of value systems that decision makers might want to consider when making a final decision, not to make the decision: the tool is informational, not prescriptive. Second, where societally-based preferences are different from expectations, making those preferences explicit can help identify areas warranting further investigation. Are the differences due to true ethical or values differences, or to communication failures that could be rectified? For example, if climate change is much more highly prioritized by the expert community than by society at large during preference elicitation (Gloria et al. 2007, Grubert in prep.), this observation encourages deeper engagement with communities aimed at reaching a shared understanding of what is under discussion and potentially a reframing of the issue (Dietz 2013). Third, applying societally-based preferences to expert-derived impact categories is not a democratic process for choosing what is important (figure 1). The use of expert-defined impact categories, apparently chosen because of their seriousness, helps mitigate several concerns about utilitarian preference functions, including the idea that people can prefer outcomes that are unattainable, immoral, or frivolous (Maier 2012, Ch. 2.1.2.1). A transparent method for assessing societal priorities for LCA impact
categories can help spur meaningful discussion and emphasize areas for improvement and further engagement.

If one concern is that WELFARES gives society too much power over decisions, a mirroring concern is that it gives society too little. The proposed separation of expert and nonexpert judgments in LCSA (Figure 1) could be considered paternalistic or community-effacing. This is particularly relevant given the increased opportunity for including community priorities using preference archetypes rather than direct interaction with people. For example, accepting expert-defined impact categories might exclude issues that are obvious and important to communities but are not research priorities. This potential exclusion is likely to be higher risk for social impact categories given their deep embedding in community-specific cultural context and relatively lower level of maturity within the life cycle fields. If WELFARES succeeds in contributing to higher transparency about LCSA in general, however, such exclusions will at least be more clear and available for discussion.

A potentially more serious problem is that the availability of WELFARES preference archetypes could reduce perceived need for direct community engagement. Though WELFARES improves incorporation of community priorities in situations where analysts would not otherwise have performed any community engagement, such as desktop analyses, there is a risk that decision makers in other situations might forgo direct engagement in favor of the sensitivity analysis. As WELFARES-based analysis is intended to provide background information that can inform more targeted community engagement, for example by identifying trade-off hot spots and soliciting more specific feedback in those areas, practitioners should be careful to emphasize that such analysis supplements rather than replaces community engagement in decision processes.

Another issue, and one raised by several participants in the broader elicitation study, is that WELFARES data could enable better propaganda. Namely, understanding community priorities could enable actors to emphasize characteristics likely to be positively received and deemphasize characteristics a community opposes, rather than changing a project. Though such manipulation is a concern, the concern is likely less severe than for existing processes due to the added transparency of an explicit framework. Further, as a collection of archetypes intended to convey the range of value systems present in society, using standard archetypes is unlikely to present an obvious opportunity for manipulation, as the different archetypes reflect differentiated priorities.

Concerns about both over- and under-privileging society-based preference information are legitimate. Users of any preference framework should be attentive to these problems, especially given that decisions using an LCA or LCSA framework are likely to include some points of contention. Case studies focused on the role of preference data in LCSA-based decision making can help test the seriousness of these concerns. Future work will present case studies based in US and Australian fossil and renewable energy communities (Grubert in prep.).

Conclusions

The LCA community faces a major interdisciplinary opportunity as it develops LCSA, a triple bottom line decision support tool that draws on multiple intellectual traditions. This article argues that LCSA’s strength as a decision support tool can be improved by introducing and standardizing intercategory preference archetypes drawn from society. Such an effort provides a means for practitioners to rapidly assess the robustness of a decision given the multiplicity of perspectives present in society, thus providing a tool for testing one of the most sensitive parameters in life cycle-grounded decision making: human values.
<heading level 1> Acknowledgements
The author gratefully acknowledges the useful feedback of three anonymous reviewers. Thanks to the Mechanical Turk Workers and Prolific Academic respondents who participated in this research both during pilots and implementation, to Rob Semmens, and to the Dynamo community. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-114747 and by an E-IPER Summer Graduate Research Grant. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. This research was performed under Stanford University Institutional Review Board Protocol IRB-33232. Stanford’s IRB can be reached at humansubjects.stanford.edu.

<heading level 1> References


### Tables

**Table 1.** Characteristics of a location for which priority data are elicited via pairwise comparisons for the proof of concept study

<table>
<thead>
<tr>
<th>Environmental characteristics</th>
<th>Socioeconomic characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air quality</td>
<td>Wages</td>
</tr>
<tr>
<td>Land quality</td>
<td>Community future</td>
</tr>
<tr>
<td>Water quality</td>
<td>Sense of community</td>
</tr>
<tr>
<td>Water availability</td>
<td>Housing and transportation</td>
</tr>
<tr>
<td>Human health</td>
<td>Diversity</td>
</tr>
<tr>
<td>Ecosystem health</td>
<td>Safety</td>
</tr>
<tr>
<td>Weather</td>
<td>Local character</td>
</tr>
</tbody>
</table>
Table 2. Aggregated and cluster preference weighting values for proof of concept study characteristics (values ≥ 0.1 shown in green, <0.1 and ≥0.05 in yellow, <0.05 in red; equal weighting of all characteristics would result in values of 0.07 for each)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Aggregate</th>
<th>Archetype 1</th>
<th>Archetype 2</th>
<th>Archetype 3</th>
<th>Archetype 4</th>
<th>Archetype 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety</td>
<td>0.14</td>
<td>0.22</td>
<td>0.10</td>
<td>0.13</td>
<td>0.12</td>
<td>0.07</td>
</tr>
<tr>
<td>Human health</td>
<td>0.13</td>
<td>0.15</td>
<td>0.10</td>
<td>0.15</td>
<td>0.10</td>
<td>0.13</td>
</tr>
<tr>
<td>Water quality</td>
<td>0.11</td>
<td>0.11</td>
<td>0.17</td>
<td>0.12</td>
<td>0.08</td>
<td>0.12</td>
</tr>
<tr>
<td>Water availability</td>
<td>0.10</td>
<td>0.12</td>
<td>0.17</td>
<td>0.10</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Air quality</td>
<td>0.10</td>
<td>0.08</td>
<td>0.10</td>
<td>0.17</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Wages</td>
<td>0.08</td>
<td>0.05</td>
<td>0.06</td>
<td>0.04</td>
<td>0.18</td>
<td>0.07</td>
</tr>
<tr>
<td>Housing and transportation</td>
<td>0.07</td>
<td>0.05</td>
<td>0.07</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Ecosystem health</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.09</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>Land quality</td>
<td>0.05</td>
<td>0.05</td>
<td>0.02</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Weather</td>
<td>0.04</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Community future</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Sense of community</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Local character</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.02</td>
<td>0.03</td>
<td>0.05</td>
<td>0.01</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>n complete datasets</td>
<td>188</td>
<td>47</td>
<td>35</td>
<td>25</td>
<td>51</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Aggregate</td>
<td>Archetype 1</td>
<td>Archetype 2</td>
<td>Archetype 3</td>
<td>Archetype 4</td>
<td>Archetype 5</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Safety</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Human health</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Water quality</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Water availability</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Air quality</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Wages</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Housing and transport</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Ecosystem health</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Land quality</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>7</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Weather</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>11</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Community future</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Sense of community</td>
<td>12</td>
<td>14</td>
<td>12</td>
<td>10</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Local character</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>12</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Diversity</td>
<td>14</td>
<td>11</td>
<td>9</td>
<td>14</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Spearman's rank correlation coefficient versus aggregate</td>
<td>0.93</td>
<td>0.85</td>
<td>0.89</td>
<td>0.75</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Percentage of sample</td>
<td></td>
<td>25%</td>
<td>19%</td>
<td>13%</td>
<td>27%</td>
<td>16%</td>
</tr>
</tbody>
</table>
Figures

**Figure 1.** LCSA is a multidisciplinary and multiparadigmatic framework, with different actors and paradigms suited best to different process stages.

<table>
<thead>
<tr>
<th>Process step</th>
<th>Essential question</th>
<th>Responsible party</th>
<th>Paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal and scope definition</td>
<td>▪ What is important?</td>
<td>▪ Expert community, with collaborative input</td>
<td>▪ Interpretivist</td>
</tr>
<tr>
<td></td>
<td>▪ How much impact is there?</td>
<td>▪ Expert community</td>
<td>▪ Positivist</td>
</tr>
<tr>
<td>Life cycle inventory</td>
<td>▪ How important are the impacts?</td>
<td>▪ Society (weightings) Expert community (selection, classification, characterization, normalization)</td>
<td>▪ Interpretivist</td>
</tr>
<tr>
<td>Impact assessment</td>
<td>▪ What will we do?</td>
<td>▪ Decision maker, with collaborative input</td>
<td>▪ Interpretivist</td>
</tr>
<tr>
<td>Interpretation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. The WELFARES proposal takes its name from common impact category themes from both environmental and socioeconomic realms.
Figure 3. Locations for participants (dots) in the proof of concept study closely match a United States population map (darker blue = higher population density).
Figure 4. Clustered archetypes ($k = 5$) show that though high priority and low priority categories are relatively consistent, preference patterns are diverse (ordered top to bottom by high to low average rating; dots represent aggregated average value for each characteristic; values for each archetype, given in table 2, sum to 100%).
Figure 5. Percentage of participants assigned to consider a parent, self, or child that express each archetype are shown as a total of the expected percentage if the distribution were even. A value of 100% (marked by the horizontal line) would mean that intergenerational assignment does not produce results different from random allocation.