

5-2016

# Whose values count: is a theory of social choice for sustainability science possible?

Mark W. Anderson

*University of Maine - Main*, mark.anderson@umit.maine.edu

mario f. teisl

*University of Maine - Main*, teisl@maine.edu

Caroline L. Noblet

*University of Maine*, caroline.noblet@maine.edu

Follow this and additional works at: [https://digitalcommons.library.umaine.edu/mitchellcenter\\_pubs](https://digitalcommons.library.umaine.edu/mitchellcenter_pubs)



Part of the [Behavioral Economics Commons](#)

## Repository Citation

Anderson, Mark W.; teisl, mario f.; and Noblet, Caroline L., "Whose values count: is a theory of social choice for sustainability science possible?" (2016). *Publications*. 35.

[https://digitalcommons.library.umaine.edu/mitchellcenter\\_pubs/35](https://digitalcommons.library.umaine.edu/mitchellcenter_pubs/35)

## **Whose Values Count: Is a Theory of Social Choice for Sustainability Science Possible**

“...most debates over social policies, decisions, and actions are fundamentally disagreements over the relevance and priority of particular values” (Leiserowitz, Kates, and Parris, 2006, p. 440).

### **Introduction: Social Choice in Sustainability Science**

If sustainability science is to mature as a discipline (Kates, 2011), it will be important for practitioners to discuss and eventually agree upon the fundamentals of the paradigm on which the new discipline is based (Gardner, 2013). The success of any emergent discipline is dependent upon clarity on the accepted elements of its paradigm. For example, Spash (2012) challenged ecological economists to explore the “philosophy of science behind ecological economics” (p. 36) and confront the difficult questions of what should be the accepted ontology, epistemology, and methodologies of that new paradigm. The same task confronts sustainability science. While it is easy to argue what from the current disciplinary landscape should not be included in sustainability science, such as benefit-cost analysis (Anderson et al., 2015), it is more difficult to agree on what should be part of the new paradigm. We offer here a continuation of this discussion, specifically an inquiry into one centrally important epistemological question to the paradigm: how do we know whose values count? The answers to this question pose a significant problem for developing a theory of social choice for sustainability science.

Sustainability is fundamentally a normative assertion (Leiserowitz et al., 2006). It is about what values will be expressed in private and public realms. It is about tradeoffs among values, both within in the current generation and between current and future generations. How society chooses the specifics among these tradeoffs, so-called social choice theory, is central to the

sustainability problem, and the very process of choosing is a choice that reflects different values (Bromley, 1998).

There is a broad literature on social choice with notable foundations from Arrow (1963) and Sen (1970), which we are not going to review here. Rather we will focus on the parts of the social choice question that are particularly relevant to sustainability science: Is social choice theory a necessary part of sustainability science? Whose values should count in making social decisions? How should the multiplicity of values that exist be known (its epistemology) and used (its methodology) in that decision process? Given the vast spatial domains and temporal domains at work in the sustainability problem, we need some means of reconciling the inevitably divergent choices depending on whose values we count, how we know what those values are, and how we count them in making social decisions. In this way we are following the admonition of Arrow (1950), that we "...must look at the entire system of values, including values about values, in seeking a truly general theory of social welfare" (p. 334).

In the first part of the paper we review the role of social choice in sustainability science and the importance of values in that social choice process. Next we discuss three temporal domains for values in social choice – the values of humans presently alive, values of those people from the past, and those of future humans. Next we review some of the thinking on non-human values. Then we discuss different ways values are weighted in current social choice processes and the problems with various current decision rules. We conclude by borrowing from Rawls (1971) and suggest a framework for deciding how to decide whose values count in sustainability science.

The objective is to continue the process of developing a coherent philosophy of science for sustainability science (Spash, 2012; Gardner, 2013; Salas-Zapata et al., 2013; Anderson et al., 2015). The contribution of this paper is to construct a framework for developing social choice theory for sustainability science. We offer an inventory of elements this theory should include and an outline of the approach to reconciling the tensions that exist among these elements. We conclude with a note of caution about the potential barriers to constructing a theory of social choice for sustainability science.

### **Sustainability Science as a Problem of Social Choice**

The breadth of writing on sustainable development, sustainability, and sustainability science is staggering. However, there is not yet coherence about the fundamental nature of these concepts. Robinson (2004) reflected one important strand of thought when he argued "...that sustainability is necessarily a political act, not a scientific concept" (p. 382). This idea of sustainability as essentially politics laid the foundation for his important work on the method of participatory backcasting (Robinson et al. 2011). However, much of the literature on sustainability science makes the claim that there is more to this than politics by another name. The claim is that there is a particular kind of science emerging. (See, for example, Kates et al. 2001; Clark and Dickson 2003; Kates 2011).

The assertion of a science of sustainability is that there are future states of society that are more sustainable than others and that these are, in part, a result of decisions we make individually and collectively today. Choice matters because, while not deterministic, sustainability is a future state of affairs that meets objectives for which there is increasing agreement in the literature. A sustainable choice for society displays four attributes (Clark and Dickson 2003; White 2013):

- Meets the needs of people in the present
- Meets the needs of people in the future
- Reduces poverty
- Conserves the planet's life support systems

These broad objectives may be met by many possible future states (see Anderson 2013, Figure 1, p. 45). Society is faced with the problem of determining what actions today will lead us toward one of these sustainable future states, choices that would contribute to a sustainability transition (Leiserowitz et al. 2006). Likewise we need to be able to identify what actions today move us away from sustainability. How society effects this transition may look like the politics Robinson identifies, but determining what needs to happen is an open question for the science of sustainability. One central challenge to developing this new type of science is the need for a social choice theory to determine just what future states meet the four objectives set for sustainability.

The call for such a social choice theory is echoed throughout the sustainability science literature, if not always using exactly the same language. For example, Jerneck et al. (2001, p. 72) describe a sequence of stages for learning about sustainability. The sequence includes 1) scientific understanding, 2) sustainability goals, 3) sustainability pathways and strategies, and 4) implementation. Stages two and three describe the key elements of social choice. These stages reflect some of the key questions Kates (2001, p. 19450) poses for sustainability science. Two of these questions describe social choice theory:

“How can society most effectively guide or manage human-environmental systems toward a sustainability transition?”

“How can the sustainability of alternative pathways of environment and development be evaluated?”

These questions challenge the sustainability science community to develop a theory of social choice, a challenge echoed throughout the literature (Kates and Parris, 2003, p. 8067; Baumgartner and Quaas 2010, p. 448; Salas-Zapata et al. 2013, p. 609; Wiek et al. 2001, Table 2, p. 213; Gowdy 2005, p. 220).

In each case the authors recognize either explicitly or implicitly that the challenge for the sustainability science paradigm is to find a way to balance competing values inherent in the needs of the present, the needs of the future, and the preservation of environmental systems. If the claims of sustainability science that it is a new paradigm are to be realized, a system of choice among different sets of values is necessary.

### **Importance of Values in Public Policy**

Almost 50 years ago John Krutilla (1967) and Garrett Hardin (1968), in two different ways, asserted the importance of values in public policy for issues we would now recognize as sustainability, even though the term was not used at that time. Krutilla argued that the concept of conservation needed to be expanded to include values “for which the market fails to make adequate provision” (p. 778). He went on to elucidate the importance of concepts economists would now recognize as option value, bequest value, and existence value (see Figure 1), suggesting that each of these entailed “...all of the problems of organizing a market for public goods...” (p. 782). For him, because humans might hold these values for natural areas, there was a potential adverse effect on the welfare of future generations. Society did not have the means of recognizing the values in the choices it made about natural area development. Furthermore,

Krutilla asserted that “homogenized” conservation land alone would not be sufficient because of the diversity of “tastes” to be accommodated. He recognized there are many different values to count and conflicts among these values need to be reconciled. Society cannot avoid making choices.

Hardin’s (1968) famous article “The Tragedy of the Commons” was at its core about this social choice issue as well. Hardin identified a category of problems for which there is “no technical solution,” where, “A technical solution may be defined as one that requires a change only in the techniques of the natural sciences, demanding little or nothing in the way of change in human values or ideas of morality” (p. 1243). In Hardin’s understanding, purely technical solutions to social problems imply we do not have to find a way to appreciate other people’s values or to devise a system of choosing among multiple values. Yet, it appears to us that increasingly sustainability problems are in the Hardin no-technical-solution category and that we will need to confront this uncomfortable fact more directly. Robinson (2004, p. 371) saw a basic division in those concerned with sustainable develop or sustainability as a split between being concerned with a “...technical fix or value change.” We need to account for values when we make social choices, which makes us confront the questions: whose values will we count and how will we count them?

To answer these questions we must know what we mean by values and whether there some commonality in how humans think about values? Schwartz (1994, p. 20) defined a value as “...a (1) belief (2) pertaining to desirable end states or modes of conduct, that (3) transcends specific situations, (4) guides selection or evaluation of behavior, people, and events, and (5) is ordered by importance relative to other values to form a system of value priorities.” This is consistent with the definition of values related to sustainability problems from Leiserowitz et al.

(2006), "...as abstract ideals that define or direct us to goals and provide standards against which the behavior of individuals and societies can be judged" (p. 418).

Schwartz (1994, Figure 1, p. 24) argued that empirical research showed a near universal structure to the values humans. While there may be a universal structure to how humans express values, there is difference among humans in which values are judged to be more important. "...(A)ctions taken in pursuit of each type of values have psychological, practical, and social consequences that may conflict or may be compatible with the pursuit of other value types" (p. 23). Economists would recognize this as a fundamental problem of tradeoffs.

For the individual, values are a central part of one's world view<sup>1</sup>, "...a combination of a person's value orientation and his or her view on how to understand the world and capabilities it offers, the lens through which the world is seen" (van Egmond and de Vries, 2011, p. 855). Individuals have different world views (ideologies), in part because they embrace different mixes of the value types Schwartz identified (Noblet et al., 2013). Different world views make it difficult to accept what Hardin would have called technical arguments about problems. Campbell and Kay (2014) showed that how individuals process technical information about a problem is affected by their world view or ideology, resulting in a "solution aversion." This is where "...solutions...are more aversive and more threatening to individuals who hold an ideology that is incompatible with or even challenged by the solution" (p. 809). Values matter even in evaluating what are seemingly purely technical issues.

One central debate in the literature about human values is the extent to which they can change. In the neoclassical economics tradition, individuals' values (tastes) are assumed to be

---

<sup>1</sup> While economists and political scientists often refer to this idea as one's world view, among social psychologists the term used more often is ideology. We will use both terms interchangeably.

constant over time and the only things that change are prices, incomes, and information to consumers (Stigler and Becker, 1977). Others, particularly those concerned with environmental values, suggest that values are not only subject to change but also need to change if we are to address sustainability problems. "... (If our values changed we would make decisions that are more protective of the biophysical environment" (Dietz et al., 2005, p. 336). In the context of sustainability, "Global value change is considered essential..." (Leiserowitz et al., 2006, p. 417). Dunlap and colleagues (2000) argued that the expression of values through environmental world views is measurable and that those world views of some populations have changed in recent decades.

The measurement of values, as revealed through preferences, is central to neo-classical economic theory (Mishan, 1981). During the last two decades of the 20<sup>th</sup> Century, which values could be known and measured expanded, partly in response to Krutilla's call to reconsider the idea of conservation. While benefit-cost analysis (BCA) had been used as a social choice methodology for water management projects for decades (Atkinson and Mourato, 2008), there grew an interest in expanding which values would be included in BCA. From this came the concept of total economic value (Randall, 1987). This placed valuation of aspects of nature firmly in the domain of welfare economics and showed the growing recognition that valuation needed to include more than just market data (Just et al., 1982). Applied economists developed new tools to measure the kinds of phenomena Krutilla had identified as central to conservation two decades earlier (Bishop, et al., 1987). There was an explosion of literature around this idea of total economic value, both in terms of what it should include and in how to measure it beyond the use of market prices as indicators of value (Champ et al., 2003; Parks and Gowdy, 2013).

After the introduction of the idea of total economic value, there was much debate over what value humans express about nature. Figure 1 (adapted from Anderson and Teisl, 2013) presents a framework for understanding the diversity of values that might be counted in decision-making for sustainability (see also Davidson, 2013). Social choice in the welfare economics paradigm had been limited to measuring human-centered values of individuals, including mostly use-based values from consumption and from non-consumptive use (I.A.1. in Figure 1). These were usually measured from market data. The innovations from non-market valuation mentioned above, allowed for the expansion of the measurement of values to include other individual values, both use-based and non-use-based, like those identified by Krutilla (I.A.2. in Figure 1). So now we can think of measuring the value of a tree, for example, as firewood (consumptive use), as a place to sit and enjoy nature (non-consumptive use), as a store of energy for future use (option value), as a gift to future humans (bequest value), or the good feeling knowing that there are trees on Earth (existence value).

Norton (2005) argued that this approach was still too narrow because it focused on individuals (“methodological individualism,” p. 203-204) and ignored values held communally, which might include value associated with both nature and culture (I.B. in Figure 1). Values of this type were also referred to as “shared and social values” which Kenter et al. (2015) described as multidimensional (p. 89).

Others have argued that there are values intrinsic to the natural world that might be called biocentric values (Nash, 1989). These are aspects of nature that are valuable without reference to humans and may focus on individual organisms (the tree as having the right to enjoy its own existence), species (the tree as part of biodiversity), or whole ecosystems (the tree as part of the function of the system) (II. in Figure 1). It is difficult to speak of intrinsic values in ways

comparable to considering anthropocentric values (Norton, 2005, pp. 161-166), but it is important to recognize that both broad types of value are potentially part of social choice processes for sustainability science.

The possibility that many different types of value exist is a challenge for sustainability science. Sustainability is often depicted as a Venn diagram (for example Parkin, 2000) where the intersection of social, environmental, and economic pillars of sustainability yields a magical sustainability area. Given the diversity of values expressed by humans, we are skeptical that there is such a sweet spot to be found. We agree with Leiserowitz et al. (2006) that, “Although nearly all participants may agree in the abstract about the importance of each of these three “pillars,” there are clearly strong tensions between these values, which often underlie the heated debates over concrete decisions and actions (p. 440).” Values matter fundamentally and thus sustainability largely is a “no technical solutions” problem. It is also a problem that has a strong temporal dimension because of sustainability’s inherent intergenerational concern (Anderson, 2013).

### **Values: Past, Present, and Future**

As an intergenerational concern different from traditional welfare models, the practice of sustainability science needs to consider values across the temporal domain. Social choice has been dominated by neoclassical economic theory of welfare economics (Just et al., 1982). In this paradigm, several assumptions are made about the temporal dimension of whose values count. Market data are the first choice for measuring values because they are deemed to be valid and reliable in that they are revealed preferences of consumers. Individuals are considered to be the best judges of what is valuable to them and their actual behavior in markets is the most reliable

indication of their underlying preferences. Furthermore, preferences are assumed to be fixed, sometimes referred to as being exogenous to the decision process and not affected by the context within which the preferences are expressed (Stigler and Becker, 1977).

Critics of this model suggested that the assumption of the “sanctity of individual choice” at the heart of the neoclassical model is flawed. “Evidence from behavioral economics, psychology, neuroscience, and evolutionary biology point to the need for an alternative to the neoclassical model of human behavior based on the fact that human values are social not individual (Parks and Gowdy, 2013, p. e8). The implication of this is that “...there may be a social interest in influencing individual preferences...” (Norton, et al. 1998, p. 194). This perspective assumes that preferences are not fixed, particularly for public policy, a perspective increasingly supported by experimental findings (Quoidbach et al., 2013; Zaval, et al., 2015; Noblet et al., 2015).

What counts in market transactions are *ex ante* values, as expressed through consumers’ preferences revealed before purchase decisions. These are understood as consumers’ willingness to pay for a good or service. We know from behavioral economics research that this value is different from the preferences that consumers would express for the same good if they already owned it and were willing to sell (Kahneman et al., 1991). Since market data cannot easily capture values other than willingness to pay before the fact, those are the values used.

Nonmarket valuation (Champ et al, 2003) was developed to estimate the equivalent to market data for values generally not captured through market transactions, like those values that Krutilla (1967) was concerned with. These include estimates of the monetary value of conserving a resource in case it might be useful in the future (option value) or just because we feel good knowing the resources is there even though we do not want to use it (existence value).

While non-market valuation broadened the realm of preferences included in social choice, these preferences were limited conceptually to the same temporal scope as market data, willingness to pay for a good or service prior to consumption (Knetsch, 1990).

For people concerned about “intergenerational moral obligations” (Norton, 2005, p. 321), this use of market data for assigning values in social choice models is one form of “presentism.” The values that we can know are those based on willingness to pay of present consumers before consumption, weighted by income. Past and future values do not count in the same way in our current approach to social choice.

This is not to say that the values of past humans are completely irrelevant in social choice models, though they are usually captured in terms of limited constraints on the current decision set from past actions. The most fundamental way this is seen is in constitutionalism, the idea that future decisions in a democracy are constrained by the rule of law codified in common law or in a written constitution. So in the United States the rules of political, social, and economic decision-making are constrained by the framework in a written constitution. It is not that the rules are unchangeable, but change is subject to interpretation by the judicial system and to rules for changing the original language that are cumbersome and rely on assent of a super majority. The impact of past values on social choice through legal mechanisms obviously varies from one political culture to another. For example, in the system in the United Kingdom, where there is no written constitution, no Parliament may bind a future Parliament in the way that the founders in the U.S. bound future Congresses through the strictures of a written constitution not easily changed.

While less constraining than written constitutions, other legal instruments are used to preserve values of the past into the future. Conservation easements and other deed restrictions constrain future land use, although what is allowed and the extent to which such property rights are honored varies from jurisdiction to jurisdiction (Morris, 2008). Even in the United States where there is robust recognition of the importance of honoring such restrictions in property records, courts may still choose to suspend these past judgments of value and override such deed restrictions. Attempts to codify bequests or legacies are always subject to the possibility that those in the future will decide it is not in their best interests to honor such attempts (Anderson, 2013). Constraints imposed by the past can be overridden by those in the present, though not always easily.

Golub et al. (2013) argue that the past also has a stake when social choice entails elements of past injustices. Such a restorative justice ethic reflects an understanding that “...intergenerational equity can reflect a concern for previous generations as much as it can for future ones” (p. 260).

Finally there is one area where the past is clearly disenfranchised in the current approach to social choice, the arena of sunk costs and benefits. In benefit cost analysis, costs that have been incurred and benefits that have been enjoyed in the past are assumed to be irrelevant to decisions about the future. The process is to start the measurement of benefits and costs in the present and then look forward, not back. This principle is so obviously accepted that it is rarely even discussed in the social choice literature. Use of past benefits and costs in decision making is simply assumed to be irrational behavior in neoclassical theory.

Thinking about the future is treated differently from thinking about the past, at least to the extent that the future is relevant to those alive in the present who will live at least some time into the future. Social choice in current practice does look forward, but in a very limited way. The preferences of the future, if anything, are assumed to be the same as those of the present measured by the *ex ante* willingness to pay of present consumers. In neoclassical economics, "...the tastes, the preferences, of future generations are something that we don't know about. ...all we can do in this respect is to imagine people in the future being much like ourselves..." (Solow, 1993, p. 191). Even critics of the neoclassical model take this approach. For example Bromley (1998) said that "We cannot know future tastes and preferences..." (p. 234), but for him this means that we cannot assume that future preferences will be like the preferences of the present. Norton (2005, p. 322) called this the "ignorance of values problem."

Others have argued to the contrary that there are means for the present to know more about the values of the future. Vergrat and Quist (2011) suggested that backcasting is one way to do this. In this approach the task of those in the present is to imagine a desirable future and working back to the present to identify the path from now to that desirable future. This is akin to Bromley's (1998) concept of a "...constructed order defining a socially acceptable program now into the future." Robinson et al. (2011) argued that this backcasting should be participatory, which gives the process an important "social learning element." In another approach, Anderson et al. (2012) proposed retrospective assessment where people in the present judge past decisions as a model for developing insights into how the future would respond to the actions of the present, how their preferences might change. In this vein of the future as stakeholders, a recent court case in the U.S. Commonwealth of Massachusetts on the divestment of stocks in fossil fuel companies by college endowment funds explicitly asserts that future generations should have standing in

some public policy issues, whether or not we are clear on the exact nature of their preferences (Harvard Climate Justice Coalition v. Presidents and Fellows of Harvard College).

Present thinking about the future is not a function of unchanging preferences. Empirical evidence shows that retrospective assessment does change how people respond to public policy proposals about future resource allocations (Noblet et al., 2015). Moreover a growing body of evidence shows that humans underestimate how much their values are likely to change in the future (Quoidbach et al., 2013). This suggests more broadly that current values are poor proxies for future values, whether the values of those currently alive who will live into the future or the values of those humans who are not yet alive.

Whether the preferences of the future are assumed to reflect those of the present or they are different in some meaningful way, an additional question is how far into the future we are going to peer in making our choices. Norton (2005, p. 321) calls this the “distance problem.” Were we to include the future in social choice models, how many generations into the future to consider would be an open question.

### **Non-Human Values in Decisionmaking**

Perhaps even more problematic than determining which anthropocentric values are to be counted is the question of non-human values. The literature on biocentric or intrinsic values of nature is vast and rich, and it is well documented elsewhere (Palmer et al., 2014; Norton, 2005, Chapter 2). For our purposes, one central question in social choice for sustainability science is: what non-anthropocentric values should be used in sustainability policy? In many ways, sustainability is a human-centered idea. In the classic Brundtland definition, sustainable development is “...development that meets the needs of the present without compromising the

ability of future generations to meet their own needs” (WCED, 1987). The focus on meeting intra- and inter-generational human needs implicitly includes instrumental values of aspects of the natural environment (I. in Figure 1). Even existence value, which is often confused with intrinsic value, is clearly human centered. This is the value humans get from knowing that some aspect of nature exists whether or not they care to use that aspect, study it, worship it, or make it available to future generations.

More difficult in sustainability science is whether intrinsic values of nature should be considered in making social choices. By intrinsic value we mean value of aspects of nature in and of itself, without reference to humans, one of several different meanings that environmental ethicists ascribe to the term (Palmer, et al. 2014, p. 422). Even if sustainability scientists opt for using these values in social choice, there remains the question of in what way these values should be considered. Is it the individual organism, a species, or the whole ecosystem that is the unit of analysis for value? (See Section II in Figure 1; see also Palmer, et al., 2014, pp. 426-429.)

Biocentric values remain controversial, even among environmentalists. Norton (2005, p. 180-190) argued from the perspective of environmental pragmatism that thinking about a dichotomy between anthropocentric values, economism in his terminology, and intrinsic value theory is misplaced. Pragmatists argue that both approaches lead to similar, if not identical outcomes in policy, so it is better to avoid a sharp theoretical distinction in types of values and focus on outcomes instead. Others argued that humans have so changed nature, domesticated it, that the focus should be on the ways in which we can manage domesticated nature to best provide for human welfare (Kareiva et al., 2007). Implicit in this argument is that domesticated nature has lost its intrinsic value and need not be considered in social choice processes.

Thus, at least four broad categories of values might be included in social choice processes in sustainability science. These are: values of humans past, present, and future, and values associated with nature without reference to humans. Even after deciding which values should be included in social choice for sustainability science, two problems remain. First is whether or how to weight these values in a decision process, and second is what decision rules should be applied in the use of diverse, and probably conflicting, values. The weighting problem is not new, it has long been recognized in the social choice literature.

### **Weighting Values for Social Choice in Sustainability Science**

It not only matters who values count, but also matters how the values used in decision making are weighted. The paradigm of neoclassical welfare economics has become the dominant paradigm for social choice in the past fifty years. The paradigm is operationalized through benefit cost analysis (BCA), although this is not without critics (Padilla, 2002; Gowdy, 2005; Parks and Gowdy, 2013; Anderson et al., 2015). Different values are weighted in the methodology, sometimes explicitly, other times implicitly.

First of all, values are weighted in terms of measurability, what Hardin (1968, p. 1244) called the commensurability issue. In current practice if a value can be expressed in monetary terms it is included; if it cannot be monetized it becomes at best part of informal political processes around the use of BCA (Atkinson and Muorato, 2008). By definition, this leaves intrinsic values of nature out of consideration. These values certainly cannot be captured in monetary terms and are likely not to be measurable in any meaningful sense. The very act of measurement probably moves them from the realm of intrinsic value to that of anthropocentric values.

Values are weighted using incomes of individuals in BCA, with the values of higher-income individuals getting higher weights and the values of lower-income individuals, of which there are more, getting lower weights (Just et al. 1982). This is because the method uses market prices to value traded good or market price analogs for non-market goods and services. Market prices assume the current distribution of income in assigning monetary values to goods and services; those with higher incomes get more “votes” in the marketplace. Techniques for non-market valuation also give more weight to the values (preferences) of higher income individuals because these techniques were designed to be analogous to market valuation.

Values of the past are completely excluded from this analysis under the sunk cost dictum: “Economic theory implies that only incremental costs and benefits *should* affect decision. Historical costs should be irrelevant” (Thaler, 1980, p. 47). This is notwithstanding experimental evidence for a sunk cost effect in individual human behaviors; that is people recognize sunk costs in making personal decisions (Arkes and Blumer, 1985). The welfare economics framework assumes that rational economic behavior is to only use present and future costs and benefits for social choice. Past cost and benefits are assumed to be irrelevant to social choice processes, even those costs incurred or benefits enjoyed just yesterday. Another way to state this is to say that in social choice zero weight is given to values of the past. Holding to this rigidly would mean recalculating social choice models repeatedly because yesterday’s costs and benefits are no longer relevant. This codifies the present point in time bias in the BCA methodology. As Mishan (1981, p. 507) said, “...the outcome of a cost benefit analysis will depend, in general, upon the reference point in time chosen. Since no purely economic consideration favors one point of time over another, the choice of any one point of time, such as the present, is arbitrary.”

Sunk costs and benefits still play an indirect and limited role in current social choices processes. They provide information for both individual and social learning. Individual preferences for future consumption are shaped by past consumption experiences of the same or similar goods or services. While this has some effect at the personal level, historians have argued that there has not been similar use of history, which would include understandings of sunk costs and benefits, in collective decision processes for public policy (Guldi and Armitage, 2014).

Current social choice models treat the values of the future differently than the past. The values of the future, therefore the monetary values of costs and benefits, are assumed to be like those of the present as we discussed above. However, those future costs and benefits are discounted (see Padilla, 2002; Atkinson and Mourato, 2008). That is, they are given less weight the further one projects into the future. These two ways of treating the values of the future both reflect the phenomenon of presentism in the current approach to social choice (Norton, 2005, p. 321).

Since different values clearly may be in conflict with each other (Schwartz 1994, p. 23), how we weight values clearly will matter in social choice. Perhaps even more important are the decision rules applied once various types of values are weighted.

### **Decision Rules for Social Choice in Sustainability Science**

Regardless which values are chosen for making decisions and how those various values are weighted, to make social choices there must be some rule for making public policy decisions about sustainability. Social choice theory for sustainability science requires a decision rule to evaluate competing policy options. Does a given policy contribute to a sustainability transition or not? The first question in terms of decision rules is that of the unit of analysis. In current social

choice models, like neoclassical welfare economics, decision rules assume that some accounting of individual values can be used to understand what is best for the collective. Norton (2005) called this methodological individualism. “Applied to normative analysis, the principle of methodological individualism implies that all benefits or goods must be understood as benefits or goods of individuals” (p. 238). He argued that some relevant values might be held and expressed communally rather than just individually (Figure 1 segment I.B.).

In the currently dominant paradigm, the decision rules use individual values in a benefit cost analysis (BC) framework. The BCA approach uses a utilitarian perspective to do this (Rawls, 1971, pp. 22-27). The approach asserts that a choice is unequivocally desirable as long as some people are made better off and no one is made worse off, a so-called Pareto improvement (Just et al., 1982, chapter 2). Since most choices impose costs on some people and generate benefits for others, Pareto improvements are not common. To solve this problem, a compensated Pareto improvement is one where those who gain welfare from a policy choice are able to compensate losers for their costs and still experience a net improvement in welfare. Since such compensation is difficult to accomplish in the real world, welfare economics adopted the so-called Kaldor-Hicks rule (Mishan, 1981, pp. 302-312; Farrow 1998). Under this construct, society is assumed to be better off from a policy choice if winners could compensate losers even when such compensation is not made. This can be thought of as a “sum of utilities” decision rule for social choice (Arrow, 1973).

In the welfare economics framework as traditionally applied, the market was assumed to achieve this utilitarian outcome as long as market failures were addressed (get the “right” prices to measure costs and benefits) and the future’s interests are assumed to be captured by present measures of utility. Bromley (1998) argued that this assumption of a “spontaneous order” fails to

meet the expectations inherent to sustainability for what the future might expect from us. His alternative is a “constructed order” in an “overlapping generations model.” “The idea of rights-based sustainability recognizes that provisioning must occur within, that is constrained by, a prior specification of institutions that will protect the interests of future persons. The notion of social bequests, protected by a regency, is consistent with this concept of rights-based sustainability” (p. 239). Bromley’s constructed order admonishes us to consider the interests of the future, but does not provide the concrete decision rules on how to accomplish this when there is the inevitable conflict among the values of the diverse parties interested in these processes, what Norton (2005, p. 200) termed moral pluralism.

Early critics of the neoclassical economics decision rules proposed multi-stage decision models as ways to solve some of the problems identified in the neoclassical model. Page (1977, p 203 ff.) suggested that conflicts between criteria for conservation and economic efficiency could be reconciled by apply the conservation criterion first (a guarantee of “permanent livability”). This would then set the constraints within which economic efficiency criteria could be applied. This approach by Page was extended by Norton (1989) when he developed criteria for thinking about what should be conserved first. Padilla (2002, p. 79) developed this even further with a multi-step evaluation process that explicitly filters types of decisions for their intergenerational and sustainability effects.

An alternative to market-based decision rules is to think about “democracy” as the procedure for reconciling value differences, particularly those between generations. “It is possible to retain a commitment to democracy and to discuss the appropriateness of values because the democratic commitment is mainly procedural, while assertions of appropriateness are put forward as empirically and morally supportable theses regarding what our obligations to the

future are” (Norton et al., 1998, p. 200). This is another way of thinking of discourse as a possible means of reducing community conflict over values through “deliberative rationality” (Deitz et al., 2005, p. 363-364). A natural extension of this was the idea of “participatory valuation” (Parks and Gowdy, 2013, p. e8). Particularly challenging is how to include the future in such deliberative democratic processes (Ekeli, 2005). Again, though, this falls short of providing a coherent set of decision rules for choosing among conflicting values once we have chosen which values are to be counted. Democratic processes with majoritarian decision rules can have similar effects to utilitarianism in terms of how values are weighted. These approaches are more consistent with the idea of sustainability as another form of politics (Robinson, 2002).

Perhaps most promising for dealing with diversity of values is multi-criteria decision making (Wang, et al., 2009), where decision makers must specify the values to be counted, the weights to be used, and the decision rules to be applied. One important aspect of this approach is that it forces explicit consideration of the decision process’s elements and allows sensitivity analysis of different decisions about what to count and how to count them. But that still leaves open the question of what process should be adopt for making these decisions about how to decide? Rawls’ theory of justice offers a way to think about this problem.

### **An Approach to Deciding How to Decide**

The fundamental problem in deciding how to decide is that of self interest. Given the temporal and spatial dimensions of sustainability and the diversity of values types described above, there is a challenge in choosing social choice models that do not serve the interests first of the choosers. It is natural for humans to focus on aspects of our lives that are closer both spatially and temporally. But in choosing how to choose, it would be fair if the process were to

transcend this self interest. Rawls (1971), in his seminal theory of justice, argued for beginning with an “original position” such “... that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles” (p. 18). The problem of choosing the principles that would lead to justice is analogous to the problem of choosing whose values should count and how they should be counted for sustainability science.

Rawls proposed for his theory of justice that those in this original position would be behind a “veil of ignorance.” Among other things, they would not know their particular class, natural abilities, intelligence, risk aversion, cultural context, or even which generation they belong to. “These broader restrictions on knowledge are appropriate in part because questions of social justice arise between generations as well as within them... They must choose principles the consequences of which they are prepared to live with whatever generation they turn out to belong to” (p. 137). This Rawlsian approach is directly applicable here.

Imagine we were ignorant of our personal circumstances, our gender, nationality, income, intelligence, beauty, health, or generation we lived in. Imagine further that we did not know our particular world view (ideology) and the values that underlie that world view. All we know is what Rawls proposed for his original position. “...they know the general facts about human society. They understand political affairs and the principles of economic theory; they know the basis of social organization and the laws of human psychology” (p. 137).

Given this as a starting point, we can imagine how rational people might structure a system of social choice. Values of “the future” would not be discounted, because we would not know what the future is -- we are ignorant of the generation we are in. Likewise, we would not ignore the values of the past; there are no sunk costs if we do not know our generation. We would avoid

the bias of presentism by not knowing what is, in fact, the present. From the original position it would be rational to not weight values by wealth or income because we would not know our own personal economic circumstances.

In the most radical version of the veil of ignorance (not one Rawls adopted), we would not know our species. While it is hard to imagine other species as part of the process for choosing the principles upon which to build social choice processes (Nagel 1974), it is one way of acknowledging the legitimacy of intrinsic values of nature.

Furthermore, under the veil of ignorance we might argue for decision rules that protect the values of minority perspectives from absolute dominance by majorities. The problem of minority perspectives has long plagued political theorists. In defending the proposed new U.S. Constitution, Madison (1961) in Federalist paper No. 10 addressed the problem of “faction,” a group with a particular world view in our terms. The fear even in democracy was “...that the public good is disregarded in the conflicts of rival parties, and that measures are too often decided, not according to the rules of justice and the rights of the minor party, but by the superior force of an interested and overbearing majority (Madison, 1961, p. 77). For him the recognition and protection of minority interest was achieved in the structure of a republican form of government designed in the U.S. Constitution. In a theory of social choice these protections need include both how decisions are made and what issues or things are so important that they are set aside from decision processes before they begin (Page 1977; Norton, 2005).

This approach based on Rawls’ original position does not get us all the way to a coherent framework for social choice in sustainability science. It does not provide us with an obvious practical alternative to BCA. It leaves more clarity on what we would not do than on what we

should on in making social choices for sustainability. What it does is help us understand the ways that this new paradigm might offer meaningful alternatives for addressing social problems in light of the diversity of human values. It establishes the principles for evaluating any proposed social choice frameworks that purport to lead to sustainable outcomes.

## **Conclusions**

Achieving sustainability is a problem of choice in a complex social system nested in even more complex biophysical systems. The choices we make about the economic, social, and natural aspects of human existence start with values. Indeed, the very definition of sustainable outcomes is a normative judgment based on what values count and how they are counted.

This is first a problem of epistemology for sustainability science. How can we know the range of relevant values that should be considered? Beyond this is the greater challenge of choosing a framework for social choice once we have decided what values will count. This is in itself a value-laden problem in that the weighting of inevitably conflicting values and selection of decision rules are both values-based. Sustainability science is a long way from offering a paradigm that addresses these epistemological and methodological problems adequately.

The starting point for defining the paradigm for social choice will be to remove self-interest from the process of choosing the choice mechanism. To accomplish this, a Rawlsian approach, where we think about the problem as if we were behind a “veil of ignorance,” is the starting point for removing self interest from social choice theory and working toward practical methodology for making social choices.

That said, there is one last outcome to consider, the potential impossibility of a theory of social choice for sustainability science. While we can easily identify social choice processes that

will not satisfy the criteria that we meet the needs of the present without jeopardizing the ability of the future to meet its own needs, that does not necessarily mean that it is possible to construct a social choice process that will assure this. Just as Arrow (1950) identified the impossibility of constructing a social welfare function in the welfare economics framework, a similar impossibility theory may well apply here. If we are to avoid this outcome, sustainability science must confront the problem of values and recognize the difficulty we face in reconciling the competing claims on our social and natural systems. Any practical framework for making social choices that lead to more sustainable futures must address these challenges.

**Acknowledgements:** This research was conducted as part of Maine’s Sustainability Solutions Initiative, supported by National Science Foundation award EPS-0904155 to Maine EPSCoR at the University of Maine. We are grateful for insightful comments of two anonymous reviewers who challenged our thinking for this article.

#### References

- Anderson MW (2013) Intergenerational bargains: Negotiating our debts to the past and our obligations to the future. *Futures*, 54: 43-52
- Anderson MW, Teisl M. (2013) Values In Anderson, R. et al. (eds), *Berkshire Encyclopedia of Sustainability: The Future of Sustainability*. Great Barrington, MA: Berkshire Publishing: 211-218
- Anderson MW, Teisl M, Noblet C (2012) Giving voice to the future in sustainability: Retrospective assessment to learn prospective stakeholder engagement. *Ecological Economics* 84:1-6
- Anderson M, Teisl M, Noblet C, Klein S (2015) The incompatibility of benefit–cost analysis with sustainability science. *Sustainability Science*, 10(1), 33-41.
- Arkes HR, Blumer C (1985) The psychology of sunk cost. *Organizational behavior and human decision processes*. 35(1):124-140

- Arrow KJ (1950) A Difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4): 328-346
- Arrow KJ (1963) *Social Choice and Individual Values*. Yale University Press, New Haven
- Arrow KJ (1973) Some ordinalist-utilitarian notes on Rawls's theory of justice. *Journal of Philosophy* 70(9):245-263
- Atkinson G, Mourato S (2008) Environmental cost-benefit analysis. *Annual Review of Environment and Resources*, 33: 317-344
- Baumgärtner S & Quaas M (2010) What is sustainability economics? *Ecological Economics*, 69(3): 445-450
- Bishop RC, Boyle KJ, & Welsh MP (1987) Toward total economic valuation of Great Lakes fishery resources. *Transactions of the American Fisheries Society*, 116(3):339-345
- Bromley DW (1998) Searching for sustainability: the poverty of spontaneous order. *Ecological Economics*, 24(2): 231-240
- Campbell TH, Kay AC (2014) Solution aversion: On the relation between ideology and motivated disbelief. *Journal of personality and social psychology*, 107(5):809-824
- Champ PA, Boyle KJ, & Brown TC (Eds.) (2003) *A Primer on Nonmarket Valuation* (Vol. 3). Springer
- Clark W C, & Dickson N M (2003) Sustainability science: the emerging research program, *Proceedings of the national academy of sciences*, 100(14): 8059-8061
- Davidson MD (2013) On the relation between ecosystem services, intrinsic value, existence value and economic valuation, *Ecological Economics*, 95: 171-177
- Dietz T, Fitzgerald A, Shwom R (2005) Environmental values. *Annu. Rev. Environ. Resour.*, 30: 335-372
- Dunlap RE, Van Liere KD, Mertig AG, Jones RE (2000) New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale *Journal of social issues*, 56(3):425-442
- Ekeli KS (2005) Giving a voice to posterity—deliberative democracy and representation of future people, *Journal of Agricultural and Environmental Ethics*, 18(5):429-450
- Farrow S (1998) Environmental equity and sustainability: rejecting the Kaldor-Hicks criteria. *Ecological Economics*, 27(2): 183-188
- Gardner SK (2013) Paradigmatic differences, power, and status: a qualitative investigation of faculty in one interdisciplinary research collaboration on sustainability science, *Sustainability science*, 8(2): 241-252
- Golub A Mahoney M & Harlow J (2013) Sustainability and intergenerational equity: do past injustices matter? *Sustainability science*, 8(2): 269-277

- Gowdy J (2005) Toward a new welfare economics for sustainability, *Ecological Economics*, 53(2): 211-222
- Guldi J, Armitage D (2014) *The History Manifesto*. Cambridge University Press
- Hardin G (1968) The tragedy of the commons, *Science*, 162(3859): 1243-1248
- Jerneck A, Olsson L, Ness B, Anderberg S, Baier M, Clark, E, ... & Persson . (2011) Structuring sustainability science, *Sustainability Science*, 6(1): 69-82
- Just RE, Hueth D, Schmitz A (1982) *Applied Welfare Economics and Public Policy*, Prentice Hall, Englewood Cliffs, NJ
- Kahneman D, Knetsch JL, Thaler RH (1991) Anomalies: The endowment effect, loss aversion, and status quo bias, *The journal of economic perspectives*, 5(1): 193-206
- Kareiva P, Watts S, McDonald R, Boucher T (2007) Domesticated nature: shaping landscapes and ecosystems for human welfare, *Science*, 316 (5833): 1866-1869
- Kates R, Clark WC, Corell R, Hall J M, Jaeger C, Lowe I, McCarthy JJ, ... & Svedin U (2001) Sustainability, *Science*, 292; 641-642
- Kates R W & Parris T M (2003) Long-term trends and a sustainability transition, *Proceedings of the National Academy of Sciences*, 100(14): 8062-8067
- Kates RW (2011) What kind of a science is sustainability science? *Proceedings of the National Academy of Sciences*, 108(49):19449-19450
- Kenter JO, O'Brien L, Hockley N, Ravenscroft N, Fazey I, Irvine KN, ... Williams S (2015) What are shared and social values of ecosystems? *Ecological Economics*, 111: 86-99
- Knetsch JL (1990) Environmental policy implications of disparities between willingness to pay and compensation demanded measures of values, *Journal of Environmental Economics and Management*, 18(3): 227-237
- Krutilla JV (1967) Conservation reconsidered, *The American Economic Review*, 57(4):777-786
- Leiserowitz AA, Kates RW, Parris TM (2006) Sustainability values, attitudes, and behaviors: A review of multinational and global trends, *Annu. Rev. Environ. Resour.*, 31: 413-444
- Madison, J (1962). No. 10: Madison. *The Federalist Papers*. New American Library, New York
- Mishan EJ (1981) *Introduction to Normative Economics*, Oxford University Press, New York
- Morris AW (2008) Easing conservation? Conservation easements, public accountability and neoliberalism, *Geoforum*, 39(3): 1215-1227
- Nagel T (1974) What is it like to be a bat? *The philosophical Review*, 83(4): 435-450
- Nash RF (1989) *The Rights of Nature: A History of Environmental Ethics*, University of Wisconsin Press, Madison

- Noblet CL, Lindenfeld LA, Anderson MW (2013) Environmental Worldviews: A Point of Common Contact, or Barrier? *Sustainability*, 5(11): 4825-4842
- Noblet, CL, Anderson M., Teisl MF (2015). Thinking Past and Thinking Future: An Empirical Test of the Effects of Retrospective Assessment on Future Preferences, *Ecological Economics*, 114: 180-187
- Norton BG (1989) Intergenerational equity and environmental decisions: a model using Rawls' veil of ignorance, *Ecological Economics*, 1(2): 137-159
- Norton B (2005) *Sustainability: A Philosophy of Adaptive Ecosystem Management*, University of Chicago Press, Chicago
- Norton B, Costanza R, Bishop RC (1998) The evolution of preferences: why 'sovereign' preferences may not lead to sustainable policies and what to do about it. *Ecological Economics*, 24(2): 193-211
- Padilla E (2002) Intergenerational equity and sustainability, *Ecological Economics*, 41(1): 69-83
- Page T (1977) *Conservation and Economic Efficiency*, Johns Hopkins University Press, Baltimore
- Palmer C, McShane K, Sandler R (2014) Environmental Ethics, *Annual Review of Environment and Resources*, 39: 419-442
- Parkin S (2000, November) Sustainable development: the concept and the practical challenge. *Proceedings of the ICE-Civil Engineering*, 138(6): 3-8
- Parks S, Gowdy J (2013) What have economists learned about valuing nature? A review essay. *Ecosystem Services*, 3: e1-e10
- Quoidbach J, Gilbert DT, Wilson TD (2013) The end of history illusion, *Science*, 339(6115): 96-98
- Randall A (1987) Total economic value as a basis for policy, *Transactions of the American Fisheries Society*, 116(3): 325-335
- Rawls J (1971) *A Theory of Justice*, Belknap Press, Cambridge
- Robinson J (2004) Squaring the circle? Some thoughts on the idea of sustainable development, *Ecological Economics*, 48: 369-384.
- Robinson J, Burch S, Talwar S, O'Shea M, Walsh M (2011) Envisioning sustainability: recent progress in the use of participatory backcasting approaches for sustainability research, *Technological Forecasting and Social Change*, 78(5): 756-768
- Salas-Zapata WA, Rios-Osorio LA, Troughon-Osorio AL (2013) Typology of scientific reflections needed for sustainability science development, *Sustainability science*, 8(4): 607-612

- Schwartz SH (1994) Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, 50(4): 19-45
- Sen A (1970) *Collective Choice and Social Welfare* Holden-Day, San Francisco
- Solow R (1993) Sustainability: An Economist's Perspective, In. Stavins R (Ed.), *Economics of the Environment*, WW Norton, New York
- Spash CL (2012) New foundations for ecological economics, *Ecological Economics*, 77: 36-47
- Stigler GJ, Becker GS (1977) De gustibus non est disputandum, *The American Economic Review*, 67(2): 76-90
- Thaler R (1980) Toward a positive theory of consumer choice, *Journal of Economic Behavior & Organization*, 1(1): 39-60
- Van Egmond ND, De Vries HJM (2011) Sustainability: The search for the integral worldview, *Futures*, 43(8): 853-867
- Vergragt PJ, Quist J (2011) Backcasting for sustainability: Introduction to the special issue, *Technological forecasting and social change*, 78(5): 747-755
- Wang JJ, Jing YY, Zhang CF, Zhao JH (2009) Review on multi-criteria decision analysis aid in sustainable energy decision-making, *Renewable and Sustainable Energy Reviews*, 13(9): 2263-2278
- White M A (2013) Sustainability: I know it when I see it, *Ecological Economics*, 86: 213-217
- World Commission on Environment and Development (WCED) (1987) *Our Common Future*, Oxford University Press, New York
- Wiek A Withycombe L & Redman C L (2011) Key competencies in sustainability: a reference framework for academic program development, *Sustainability Science*, 6(2): 203-218
- Zaval L, Markowitz EM, Weber EU (2015) How Will I Be Remembered? Conserving the Environment for the Sake of One's Legacy, *Psychological science*, 26: 231-236