Received: 11 October 2019 | Returned for Revision: 11 December 2019 | Accepted: 8 July 2020

Decision Analysis

Evaluating the Application of Decision Analysis Methods in Simulated Alternatives Assessment Case Studies: Potential Benefits and Challenges of Using MCDA

Christian Beaudrie,† Charles J Corbett,‡ Thomas A Lewandowski,§ Timothy Malloy,|| and Xiaoying Zhou#*

†Technical Safety BC, Vancouver, British Columbia, Canada
‡UCLA Anderson School of Management, Los Angeles, California, USA
§Gradient, Seattle, Washington, USA
||UCLA School of Law, Los Angeles, California, USA
#California Department of Toxic Substances Control, Sacramento, California, USA

ABSTRACT

We compare how several forms of multicriteria decision analysis (MCDA) can enhance the practice of alternatives assessment (AA). We report on a workshop in which 12 practitioners from US corporations, government agencies, NGOs, and consulting organizations applied different MCDA techniques to 3 AA case studies to understand how they improved the decision process. Participants were asked to select a preferred alternative in each case using a different decision analysis approach: their unaided decision-making method, individual or lightly facilitated group multiattribute value theory (MAVT), and more extensively facilitated group structured decision making (SDM). Surveys conducted after each exercise revealed that participants were positive toward the use of formal decision-making methods for AA, reporting meaningful increases in their understanding of the trade-offs involved and their own values. Participants also reported challenges with each approach. While the MCDA techniques were reported to enhance transparency and communication, they did not consistently lead to higher satisfaction with a decision and/or outcome, and they were not more likely to be adopted within their organizations than unaided approaches. More formal decision-making methods have promise in the context of AA, but practitioners will need more guidance to use such tools successfully. Practitioners will also need to define what "success" constitutes; different approaches may be called for depending on whether the objective is increased understanding, satisfaction with the outcome, satisfaction with the process, or something else. *Integr Environ Assess Manag* 2021;17:27–41. © 2020 SETAC

Keywords: Alternatives assessment Decision analysis Multiattribute value theory Structured decision making

INTRODUCTION

Alternatives assessment (AA), also known as alternatives analysis, involves the identification, assessment, and comparative evaluation of different options for satisfying a with an aim to remove or reduce hazardous materials in consumer products (Rossi et al. 2012). It attempts to be a more rigorous and transparent methodology for comprehensive assessment and decision making, directed at avoiding substitutions where one undesirable outcome is replaced by another equally or more undesirable one. The AA is integral to implementing the "substitution" principle, which provides that it is generally better to use a less hazardous chemical or nonchemical alternative in products where feasible (Swedish Chemicals Agency 2007).

This article contains online-only Supplemental Data.
* Address correspondence to Malloy@law.ucla.edu
Published 7 July 2020 on wileyonlinelibrary.com/journal/ieam.

An AA typically considers multiple criteria, including human health impacts, environmental and ecological impacts, technical performance, and economic feasibility. Each of these include several subcriteria. For example, human health impacts may include different endpoints (e.g., carcinogenicity, endocrine disruption, and neurotoxicity), exposure routes, and populations. An AA can be challenging. The data regarding the relevant criteria are often incomplete and uncertain. The available data can be qualitative or quantitative and are usually incommensurable. Selecting among alternatives can present thorny trade-offs and value-based choices. For example, how does one choose between reducing ecological impacts and maintaining critical technical performance features of a product? Or between an endocrine disrupting ingredient and a moderately carcinogenic one? Therefore, AA is not just an objective, expert-driven process, but integrates expert knowledge and judgment (often in the face of uncertainty) with value-based judgments (also sometimes under

 DOI: 10.1002/ieam.4316

conditions of conflict and uncertainty). The balance to be struck between the role of expert knowledge and value-based judgment will vary depending upon, among other things, the specific decision context.

The AA thus presents a classic multicriteria decision scenario. Multicriteria decision analysis (MCDA) can assist policy-makers, businesses, and other stakeholders engaged in AA. An MCDA is not a single decision-making approach or method, but a family of methods grounded in differing theoretical bases designed to aid decision makers facing difficult trade-offs (Belton and Stewart 2002). Despite their differences, the diverse MCDA methods all assist the user to varying degrees in 4 dimensions (Keeney and Raiffa 1972; French 1989):

- systematic structuring and communication of data and trade-offs;
- explicit articulation and consideration of stakeholder preferences and values;
- transparency and accountability for the decision-making process;
- and creation of a framework for consensus building and collaboration in a group setting.

There is limited literature on the relative value of MCDA over unassisted decision making, particularly in the context of AA. This study aims to assess whether MCDA approaches do enhance decision making in the AA context and to identify the challenges of using these tools, as well as areas for improvement. We convened a working group of 12 representatives from government, industry, and non-governmental organizations for a 2-day workshop at the UCLA School of Law on 5–6 December, 2017. The participants used various decision-analytic techniques (before and during the workshop) and provided feedback on them through surveys and discussion.

One challenge that immediately becomes apparent in assessing whether MCDA approaches can enhance decision making in AA is the absence of an objective performance benchmark: one alternative is often not objectively better than another. Rather, alternatives may be better or worse across different considerations, the relative valuations of which are inherently subjective. Therefore, we evaluated the value of MCDA approaches along several dimensions, including the extent to which participants were satisfied with the decision or the decision approach, the extent to which the method helped improve participants' understanding of their values or the information involved, and the likelihood with which the participants would use the method. No single method emerges as clearly preferred along each of these criteria.

BACKGROUND

Alternatives assessment

An AA is used in a variety of contexts. In some cases, AA will be driven by regulatory mandates. In 2013, California's Department of Toxic Substances Control (DTSC) promulgated its Safer Consumer Products regulations (DTSC 2013), requiring manufacturers in certain circumstances to evaluate potential alternatives to determine if their products can be made safer. In the European Union, manufacturers seeking continued use of certain substances of very high concern must demonstrate through an "analysis of alternatives" that no appropriate feasible substitute is available. Regulators may perform AA themselves to control which substances are used and how. In other cases, a business or NGO may perform an AA to see whether products or processes involving substances of concern (often of emerging concern) can be reasonably replaced or eliminated. An AA may provide input to a single decision maker, such as a regulatory official or senior executive, or it may support collaborative or group decision making. The particular context may influence how the AA is performed, including which decision-making approach is used.

There are numerous frameworks for conducting AA. While the details vary, most frameworks contemplate the 4 steps listed below (Jacobs et al. 2016):

- *Identification of key criteria and metrics for evaluating alternatives.* The criteria typically fall within 5 major categories: physical chemical hazards (i.e., explosivity and flammability); human health impacts; environmental and ecological impacts; technical feasibility; and economic feasibility.
- *Identification of potential alternatives.* This could include drop-in chemical substitutes, use of alternative materials, or a product and process redesign which eliminates the need for a chemical (Tickner et al. 2015).
- *Collection and compilation of data.* This could include experimentally derived as well as predicted properties (Zheng et al. 2019).
- *Evaluation of alternatives.* This involves evaluating the performance of the alternatives relative to the baseline product and each other and selecting a preferred alternative. However, most AA frameworks are primarily descriptive; even if they include quantitative elements and risk-based ranking, most provide minimal guidance for making that ultimate decision.

Illustrative scenarios of AA in practice

An AA could take many forms in practice. Rather than describe all possible scenarios, we briefly sketch 3 different contexts in which AA might be used in practice. These are not meant as complete descriptions of actual scenarios, but as exemplars of some of the main ways in which the application context may differ. Later, when we interpret our findings from the workshop, we will relate them back to these 3 scenarios.

Scenario 1: A large company is concerned about the health impacts of a chemical used in its consumer products, even though use of the chemical is legal. An individual within the company is asked to perform an AA to inform the selection of an alternative to the substance. This individual conducts the AA independently with expert input from others within the company regarding the performance of the potential alternatives across a range of relevant criteria. The individual makes a recommendation to senior management who ultimately has to choose which alternative to proceed with. The ingredients in the products are proprietary and the AA and ultimate decision regarding alternatives will be kept confidential.

Scenario 2: A trade association tasks a working group with recommending the best alternative for a substance that is about to be banned. This group will conduct an AA to guide its consensus-based decision making. The group has representation from several diverse members of the trade association and needs to incorporate the varying values and objectives of its members in determining which alternative to propose. The trade association needs to be able to explain its choice to their members and to the public at large.

Scenario 3: A government agency, such as the federal EPA, is considering a phaseout of an existing chemical and is conducting an AA to confirm the availability of viable alternatives. The agency will not adopt a phaseout unless it is satisfied that viable alternatives are or will be in the marketplace. An agency work group will perform the AA internally, using existing data, but will invite interested stakeholders to submit relevant data regarding the performance of the potential alternatives across a range of relevant criteria. The AA will be used by senior management at the agency to support a decision based on the available alternatives. The decision and the process which led to the decision need to be documented and made available to outside stakeholders, including the public, NGOs, companies, and others, any of which may contest the decision, possibly in court. The specter of such legal proceedings clearly changes the nature of documentation required for every decision made, including how public comment was taken into account.

Decision methods

A wide range of decision-making methods is available to AA analysts, ranging from informal rules of thumb to highly complex, statistically based methodologies (NAS 2014; Malloy et al. 2017). We categorize decision-making tools and methods, defined as formal and informal aids, rules, or techniques that guide specific decisions, into 3 general types: narrative, rule-based, and analytical. Table S1 in the Supplemental Data provides an overview, and for the purpose of this paper further distinguishes between 3 different analytical approaches: multiattribute value theory (MAVT), outranking, and structured decision making (SDM). (Each of these general methods could be used by individuals or groups with varying degrees of facilitation.)

Narrative methods involve a holistic, qualitative balancing of the data and associated trade-offs. While in some cases the decision is informed by stated rules of thumb or principles, there is no quantitative scoring of the alternatives and no explicit quantitative weighting of the relative importance of the decision criteria (Linkov et al. 2004; Eason et al. 2011). The narrative approach is widely used in regulatory decision making. The federal Superfund statute and implementing regulations use it in selecting cleanup remedies for hazardous waste sites (EPA 2018). Under the name "verbal-argumentative approach" it has also been used by the European Chemicals Agency (ECHA) in prioritizing chemicals for regulation under the European Union's REACH program (ECHA 2010).

Rule-based methods, also known as elementary approaches, provide more structure and specific guidance than the narrative method, but do not require sophisticated software or specialized expertise in decision analysis. Examples include a decision tree using an ordered series of questions, a set of checklists, an integrated suite of specific decision rules, or simple algorithms, all used to frame the issues and guide the evaluation (Eason et al. 2011). Rule-based methods use both quantitative and qualitative data, and may include either implicit or explicit weighting of the decision criteria (Eason et al. 2011).

Analytic methods are tools that couple narrative evaluation with a mathematically based formal tool such as MCDA (Linkov and Moberg 2012). An MCDA consists of a variety of methods drawing on a range of theoretical bases and methodological perspectives. Some MCDA tools, such as those relying upon Multi Attribute Value Theory (MAVT), are "optimization tools" that seek to maximize the achievement of the decision maker's preferences represented as value functions (Linkov and Moberg 2012). Others, such as outranking eschew value functions, instead seek to identify "dominant" alternatives through pairwise comparisons (Malloy et al. 2013). A third type is structured decision making (SDM), a mixed approach integrating optimization and pairwise comparisons within a particularly highly structured, facilitated process, often used in group decision-making contexts with extensive stakeholder involvement (Wilson and McDaniels 2007). It uses an iterative 5 step process to clarify the decision context, identify meaningful decision objectives, develop creative alternatives that meet those objectives, estimate performance, and evaluate trade-offs, and finally to decide on a preferred alternative (Gregory and Wellman 2001; Gregory et al. 2012).

Despite their differences, these analytical methods share common features which set them apart from the basic narrative method (Belton and Stewart 2002). Each method provides a systematic, observable process in which an alternative's performance across the decision criteria is compared to other alternatives, typically in a matrix. In most, the relative importance of each decision criterion is quantified and used to weight the performance of each alternative on each criterion. This is then typically aggregated to characterize the relative overall performance of each alternative (Kiker et al. 2005). Analytical methods can also explicitly address uncertainty in the data (Malloy et al. 2013).

Effectiveness of MCDA in environmental decision making

Environmental problems in general are inherently multidimensional, and MCDA methods have been applied in this context, although not yet extensively in the context of AA (Lahdelma et al. 2000; Kiker et al. 2005; Steele et al. 2009;

Huang et al. 2011; Linkov and Moberg 2012; Cegan et al. 2017). Guitouni and Martel (1998) offer general guidelines for how to choose an appropriate MCDA method, such as to use a group rather than individual decision-making process if there are many stakeholders. Cinelli et al. (2014) compare 5 MCDA methods along dimensions such as compensation degree and ease of use; they note that in practice the specific method is usually chosen based on familiarity rather than on the actual decision context. Kain and Söderberg (2008) found substantial differences between how various MCDA methods perform on criteria such as functionality, ease of use, transparency, and trust. Others find that MCDA methods yield sometimes similar and sometimes different results on a range of environmental problems (Salminen et al. 1998; Yatsalo et al. 2007). The authors recommend using multiple methods in practice.

Several studies focus on comparing MCDA methods with unaided decision making (Arvai et al. 2001; Arvai and Gregory 2003; Wilson and Arvai 2006; Hajkowicz 2007; Bessette et al. 2014). Arvai and various colleagues (Arvai et al. 2001, Arvai and Gregory 2003; Wilson and Arvai 2006) compared unaided decision making to "value focused" decision making in a range of simulated environmental settings. In value focused decision making, participants were instructed to consider how well each alternative advanced the participant's values and objectives and to think through the trade-offs involved. This led to higher satisfaction and comfort among decision makers than unaided decision making. However, Wilson and Arvai (2006) found that although subjects claimed that their (value-focused) decisions reflected their objectives, those decisions were not consistent with their previously stated rankings of objectives. The authors caution against overinterpreting this paradox but argue that it highlights the importance of evaluating environmental decisions both in terms of process (where value-focused decision making appeared to help) and in terms of outcomes (where in their experiment it did not). We find similar apparent inconsistencies between how participants responded to the various methods they used.

Closest to our work are 2 papers by Bell and various colleagues based on a 2 day workshop for 20 climate change experts and policy-makers (Bell et al. 2001, 2003). Participants were asked to rank climate change policy options (e.g., a carbon tax of US$75 or $150, promotion of nuclear power, etc.) using their own unaided decision-making approach, then used MCDA methods and integrated assessment risk-based methods to revisit the decision. The ranking of policies varied by method (for the same person) and across participants (using the same method). Different weighting methods produced very different weights, and the participants recommended using weights from multiple methods. The participants preferred the unaided assessment to any MCDA method (Bell et al. 2003). Holistic assessment ranked highest on most evaluation criteria. Participants did not find that MCDA methods were likely to increase confidence in the decision or consistency. Almost every MCDA method was the most highly-recommended by at least 1 participant, indicating a wide diversity in preferences for methods (Bell et al. 2001). This highlights the same challenge that we encountered, associated with the lack of an objective performance benchmark for evaluating the effectiveness of MCDA methods.

There is substantially less research on the use of MCDA in AA, and none that we are aware of on the use of SDM in AA. It is now fairly accepted that MCDA methods have relevance to the AA process (Ogunseitan 2016; Zheng et al. 2019), including by the National Academy of Sciences (NAS 2014) and the California Department of Toxic Substances Control (DTSC 2017). Some research has started to explore conceptual, methodological, and normative issues associated with the use of MCDA in AA (Malloy et al. 2017; Zheng et al. 2019), and has developed criteria to guide the selection of decision-making approaches and tools, including MCDA (Malloy et al. 2015). Other work demonstrates the use of various MCDA methods in AA case studies (Zhou and Schoenung 2008; Malloy et al. 2013; Tsang et al. 2014). There has been no systematic assessment of the benefits and challenges associated with MCDA methods relative to unaided decision making from the perspective of AA practitioners. Our work begins to fill that gap. In the absence of an objective performance benchmark against which to evaluate the outcome of different AA approaches, we will not be able to point to a specific MCDA method as being superior, but we are able to identify how various methods compare along dimensions such as participant satisfaction, improved understanding, likelihood of use, and more.

METHODS

Studies of environmental decision making consistently identify 3 challenges: dealing with large numbers of decision criteria, adequately incorporating decision makers' values, and systematically addressing trade-offs (Bessette et al. 2014; Malloy et al. 2017). This study assesses the utility and limitations of various MCDA approaches in overcoming these challenges in the context of AA.

We convened a group of experts in chemical and product assessment from government, industry, and nongovernmental organizations for a 2 day workshop to test various decision-analytic techniques and provide feedback for the design of improved tools and processes. Figure 1 summarizes the structure and substance of the workshop.

Several weeks before the workshop, participants completed a simulated AA case study (detailed later), selecting a replacement for Cu-based marine paint without the use of formal decision analytic tools. The workshop began with short background presentations regarding AA, decision analysis, and the second case study, which was identical to the preworkshop case study aside from a different set of alternatives. Participants then used one type of decision software to reach a decision on case study 2 (either as individuals or part of a group) then used a different software package to reach a decision on case study 3 (in a fully

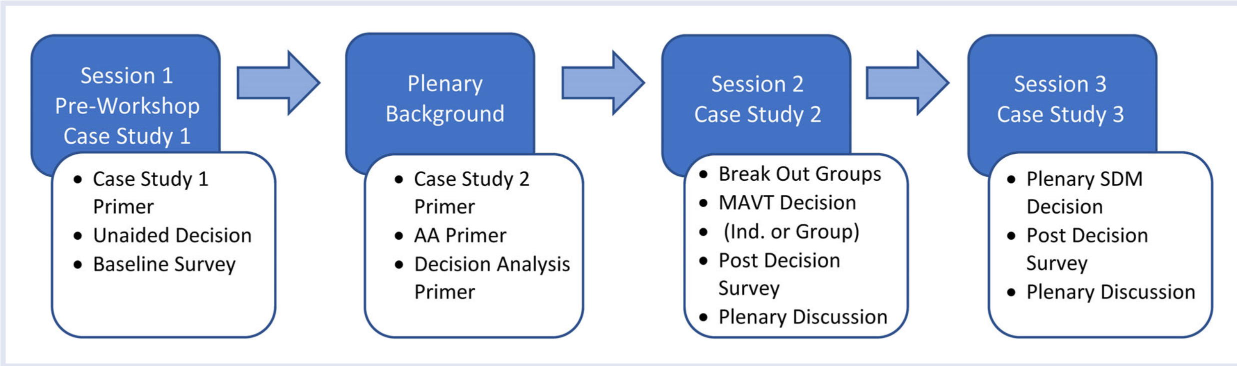

Figure 1. Workshop workflow.

facilitated plenary exercise). Further details are provided below.

After each of the 3 respective decision-making exercises for case studies 1, 2, and 3, participants individually completed an online survey regarding their experience. The surveys were designed to evaluate the performance of the alternative decision support tools against standards for "good" decision making. Given the value-based trade-offs inherent in alternatives analysis, one cannot simply ask whether a method or tool leads to the right choice. There is no correct outcome against which to evaluate or validate such methods. (Wilson and Arvai 2006). Researchers have identified a range of measures of decision quality focusing on the decision makers' self-reported experience with and perceptions of the decision process itself (Gregory et al. 2012; Bessette et al. 2014), the decision outcome (Wilson and Arvai 2006), and the likelihood of adoption of the tool or method (Hajkowicz 2007). We constructed survey questions drawing upon each of those 3 general measures. Multiple-choice questions used 5 point scales, and open-ended questions solicited additional comments (Table 1).

Participants

Due to budget and practical constraints, we targeted 12 participants. We invited approximately 20 participants familiar with the concept of AA through personal networks; some accepted, and several others suggested alternates. All participants had considerable work experience in decision making concerning product alternatives, and had at least some familiarity with sustainability, product chemistry, and toxicology, as well as regulatory compliance. We covered participants' travel, lodging, and meals; subjects received no other compensation for their time. The study was certified as exempt from review for human subjects impacts by the UCLA Institutional Review Board.

The 3 AA case studies

The case studies were built around a single semifictional narrative—the search for alternatives to Cu-based antifouling paint for recreational boats. Marine organisms can attach themselves ("fouling") to boat bottoms, possibly damaging the boat's structure and reducing its speed, maneuverability, and fuel efficiency. Boats can also carry invasive fouling organisms to other harbors. Antifouling paint applied to boat hulls can prevent or reduce the attachment of marine organisms. Many of these paints contain Cu as a biocide, an active ingredient intended to kill barnacles and other marine organisms. Copper enters the marine environment through leaching and during hull cleaning, causing marina Cu levels to exceed regulatory standards and potentially affect aquatic organisms.

The participants played the role of marine paint manufacturers responding to a government mandate to conduct an assessment of alternatives to Cu-based antifouling paint. The exercise focused on the last step of an AA—evaluation of alternatives. The case studies did not include all relevant data that would be considered in practice, nor deal with data gaps or mixtures of chemicals in products. Instead, the case studies highlighted certain thorny trade-offs presented by the alternatives.

Two types of antifouling products were considered: biocidal and nonbiocidal. Biocidal products are typically paints formulated with biocides that are released over time, containing 1 or more active ingredients such as Cu, synthetic organic biocides, or organo-Zn compounds. Nonbiocidal approaches include "photoactive" coatings that generate H_2O_2 at the boat's surface for antifouling properties and silicone-based coatings that form slippery, low-friction surfaces to which organisms have trouble attaching.

We selected a recently published report on alternatives to antifouling paint in Washington State as our starting point (Northwest Green Chemistry 2017). We constructed 3 case studies each with a different set of alternatives to Cu-based paint. We simplified the problem to 4 categories of decision criteria and related subcriteria: human health concerns, ecological concerns, technical performance, and cost. We did not attempt to structure the alternatives such that there was a "best" or "optimal" alternative; whether an alternative is preferred by a decision maker depends heavily upon the particular respective weight placed by the decision maker on each decision criterion. Rather the alternatives were designed to introduce challenging trade-offs across different criteria. We asked participants to assume that the data were sufficiently reliable and to not take data

 DOI: 10.1002/ieam.4316

Table 1. Decision metrics

	Metric	Question	Reference
Process questions	Overall satisfaction	Overall, how satisfied are you with the [decision-making] approach?	Bassette et al. 2014
	Ease of use	How would you rate the difficulty or complexity of applying [this approach] to this decision?	Lindahl 2006; Booker 2011
	Enhanced knowledge	Would you agree that using this approach improved your understanding of the available information?	Bassette et al. 2014
	Transparency	How would you rate the level of transparency that is possible using [this approach]—i.e., does it enable users to clearly articulate assumptions, decision process, and rationale?	Kain and Söderberg 2008; Gregory et al. 2012
	Understanding of own values	Would you agree that using this approach improved your understanding of your own values?	Hajkowicz 2007
	Understanding of trade-offs	Would you agree that using this approach improved your understanding of key trade-offs between alternatives?	Malloy et al. 2017
	Enhanced communication	Would you agree that using [this approach] could help you to better communicate your decision-making results?	
Outcome questions	Overall satisfaction	Overall, how satisfied are you with your decision?	Bassette et al. 2014; Wilson and Arvai 2006
	Alignment with values	Do you agree that the decision outcome using [this approach] accurately reflects what matters to you?	Bassette et al. 2014; Wilson and Arvai 2006
	Alignment with intuition	Do you agree that the decision outcome aligns with your initial impression (or gut feeling) about what is the best alternative?	Belton and Stewart 2002
Adoption questions	Overall comfort	How comfortable are you with applying [this approach] to other decisions?	
	Likelihood of adoption	How likely is it that you would use [this approach] in your institution to support chemical alternatives analysis?	Hajkowicz 2007

uncertainty into consideration. The performance matrices are set out in Table 2. More specific descriptions of the decision criteria are included in the Supplemental Data.

Decision analysis methods

Many combinations of decision methods, weighting approaches, and facilitation approaches could be used in AA. Due to the workshop time constraints we focused on only 3 methods.

Session 1: Unaided decision making. Session 1 took several weeks before the workshop convened. Prior to the workshop we gave participants a description of case study 1 with its respective performance matrix and asked them to select their preferred alternative using whatever approach they desired. Participants submitted their decision along with a narrative explanation of their decision-making method. Participants also completed a survey regarding their experience.

Session 2: Individual and small-group decision making aided by MAVT. In session 2 at the workshop, participants were split into 2 cohorts to work on case study 2 using Decerns MCDA decision support software, which is part of the framework that was initiated within the international project DECERNS (Decision Evaluation in Complex Risk Network Systems) (Yatsalo et al. 2016). Decerns MCDA was configured to implement a basic MAVT, an optimization method that seeks to maximize the achievement of the decision maker's preferences. Using value functions, the method normalizes the performance of the alternatives on each criterion to a score between 0 and 1, then multiplies that score by a weight assigned to that criterion. It aggregates the weighted scores to arrive at a total score for the alternative.

Multicriteria decision analysis approaches generally use 1 of several weighting methods: direct rating, pairwise comparison (used in analytical hierarchy process and in outranking methods), and "swing weighting" (often used with MAVT methods) (Belton and Stewart 2002; Linkov and

Table 2. Performance matrices

Decision criteria		Performance measure	Direction[a]	Potential Alternatives														
				Session 1: Unaided					Session 2: MAVT					Session 3: SDM				
				CerSmooth	SlipCote	Zn2000	Expedition	NoFoul	HullSaver	StreamXL	Barrier	Aquaslide	Armor99	BlasOt99	ThioiZ	Barnaclear	Barracuda	GuardX3
Human health concern[b]	Carcinogenicity[c]	4 pt scale	Higher	4	4	3	2	3	4	3	2	3	4	3	4	2	3	4
	Neurotoxicity (oral)[3]	mg/kg/day	Higher	50	500	250	10	0.5	0.5	700	50	75	0.5	50	10	70	Data Gap	1
	Reproductive/Develop-mental Toxicity (oral)	mg/kg/day	Higher	200	20	70	0.5	0.01	100	0.01	250	175	0.01	80	200	10	175	60
	Respiratory Allergen/Asthmogen[d]	3 pt scale	Lower	1	1	2	2	2	2	1	1	3	1	2	3	1	Data Gap	1
Ecological concern	PBTaq	100 pt scale	Lower	0	20	40	30	65	10	40	40	10	5	60	60	50	10	2
	VOCs (emissions during application)	grams/liter	Lower	150	100	300	250	270	1200	400	600	1300	200	400	100	1300	300	900
Technical performance	Longevity (time between needed applications)	Years	Higher	2	3	4	5	4	2	3	4	4	4	4	3	4	3	4
	Efficacy (performance in antifouling test)	5 pt scale	Higher	3	3	5	3	5	3	3	5	2	5	5	3	5	4	5
Cost	Cumulative 5 year cost (labor and materials)	Dollars	Lower	14,000	11,000	6,977	6,891	5,565	7,800	8,500	10,500	6,890	11,700	9,000	4,000	4,200	8,000	11,000

[a] "Direction" refers to the preferred direction of the performance measure, e.g., "higher" means that higher scores are preferred.
[b] Scored based on all ingredients present above 0.1%, worst-case ingredient drives the score.
[c] Derived from International Agency for Research on Cancer classification of carcinogenic agents: 1 = Known carcinogen, 2 = Probably carcinogen, 3 = Possible carcinogen, 4 = Evidence of noncarcinogenicity.
[d] No observed adverse effect level in a relevant animal toxicity study.

DOI: 10.1002/ieam.4316

Moberg 2012). While direct rating (i.e., simply assigning a relative weight directly to the relevant decision criteria) is typically easier for respondents to use, it may lead to significantly biased results (Németh et al. 2019). This exercise used a swing-weighting approach to elicit weights that involved showing participants the "best" and "worst" performance across alternatives for each criterion and asking them "which criterion would you first 'swing' from its worst performance to its best?" (Belton and Stewart 2002). The swing weighting approach helps participants take into account the decision context by highlighting how much the performance can be improved on each criterion when selecting between different alternatives. For example, as a general matter, a respondent might view carcinogenicity as much more important than skin irritation. But in a context in which the alternatives all performed well in terms of carcinogenicity, this criterion becomes less important for the decision since the desired level of performance will be satisfied by all alternatives. Instead, the respondent would likely view skin irritation as a more important concern for the given decision, since there is a bigger (and meaningful) difference in performance on that criterion between alternatives. Swing weighting takes this context into consideration, while direct rating typically does not.

The Cohort 1 participants engaged in individual decision making. They gathered in a small computer lab, where each participant was asked to individually rank the alternatives from most preferred to least preferred. Following this ranking exercise, 2 of the authors provided a brief tutorial on Decerns MCDA and remained available to assist participants with the software. Participants used Decerns MCDA individually to assist in the generation of criteria weights and an associated ranking of the alternatives. They each engaged in a self-guided sensitivity analysis in which they were encouraged to vary weights and vary the value function curves. After comparing the results from the ranking exercise and from their MAVT analysis, the Cohort 1 participants selected their respective preferred alternatives.

The Cohort 2 participants engaged in small-group decision making. They likewise were asked to individually rank the alternatives, after which they took part in a lightly facilitated group exercise comparing alternatives in Decerns MCDA on a shared screen. Cohort 2 participants each used an online swing-weighting tool developed by Compass Resource Management (Compass) to generate individual weights (Beaudrie and Schroeder 2017), which were aggregated with equal weight given to each participant. This was done to define weights for the group to use in Decerns MCDA, which only allows 1 set of weights at a time. The small group then reviewed and discussed the aggregate weights before entering them in Decerns MCDA to compare alternatives. They each engaged in a facilitated sensitivity analysis of the Decerns MCDA results, and after group deliberation selected their respective preferred alternatives. At the completion of the exercise, participants completed a survey regarding their experience.

Session 3: Group decision making aided by SDM. For case study 3, all participants engaged in 1 fully facilitated group decision making using a SDM method that included pairwise comparisons of alternatives (directly comparing performance in a consequence table to identify dominant alternatives and insensitive criteria, and to compare overall performance), discussion on trade-offs, elimination of poorly performing alternatives, and a MAVT exercise to assess the overall utility of the remaining alternatives. After providing a brief introduction to SDM, 1 of the authors led an extensively facilitated group session to discuss the performance of the alternatives across each objective. Participants looked for opportunities to simplify the performance matrix by removing "dominated" alternatives (those that are outperformed by 1 or more alternatives), and "insensitive" objectives and criteria (that do not vary much across alternatives). Trade-offs among objectives were then explored using pairwise comparison, and the group discussed preferences. Weights were then elicited using an online SMARTER weighting tool (developed by Compass, see Beaudrie and Schroeder 2017), and visualization tools were used to review similarities and differences in participant weights and to investigate the sensitivity of MAVT value scores to different weights. "SMARTER" is a variant of swing-weighting that simplifies the elicitation process by asking participants only to rank order the criteria. Weights are assigned using a formula based on an empirically validated association between weights and ranks (Edwards and Barron 1994). Lastly, the group debated the results and collectively arrived at a consensus preferred alternative.

RESULTS AND ANALYSIS

Preworkshop unaided decision-making exercise

Of the 12 participants who completed the preworkshop exercise, 8 chose the paint identified as "Zn2000" as the best alternative, while another selected Zn2000 as 1 of the top 2. Zn2000 scored very well on technical performance and fairly well on cost but presented concerns regarding human and ecological toxicity. We asked the participants to describe how they arrived at their decisions. We categorized their decision-making approach based on whether it was primarily narrative (4 participants), rule-based (3 participants), a mix of narrative and rule-based (4 participants), or similar in spirit to MCDA (1 participant) (see Table S2 in the Supplemental Data for more detailed descriptions of their respective decision approaches). Table 3 (survey results) shows that participants were reasonably satisfied with their unaided decision approaches (mean of 3.9 out of 5) and somewhat satisfied with their respective choice of paint alternative (mean of 3.6 out of 5) (see Figure S1 in the Supplemental Data for histograms of the results for each survey question).

Session 2: Individual and group decision making aided by MAVT

The participants in Cohort 1, the individual MAVT treatment, were quickly comfortable with Decerns MCDA and

Table 3. Summary of results

Metric	Scale		UnAided	MAVT Individual	MAVT Small Group	SDM Large Group
	Best	Worst				
1. Improved understanding of your own values	5-Strongly Agree	1-Strongly Disagree	3.3	3.8	3.9	3.8
2. Improved understanding of available information	5-Strongly Agree	1-Strongly Disagree	3.3	3.4	3.6	3.6
3. Improved understanding of the trade-offs btw alternatives	5-Strongly Agree	1-Strongly Disagree	3.3	4.4	3.9	4.0
4. Enables or promotes transparency	5-Strongly Agree	1-Strongly Disagree	3.4	3.8	2.7	3.8
5. Could help you communicate results and decision rationale	5-Very Comfortable	1-Very Uncomfortable	3.7	4.0	2.9	4.0
6. Difficulty of applying decision-making approach	5-Very Easy	1-Very diffiult	3.5	4.0	3.3	3.8
7. Difficulty of applying weighting method	5-Strongly Agree	1-Strongly Disagree	3.0	4.2	3.0	4.1
8. Top alternative aligns with your intuition or gut	5-Strongly Agree	1-Strongly Disagree	2.9	3.4	1.6	3.5
9. Top alternative reflects what matters to you	5-Strongly Agree	1-Strongly Disagree	3.9	3.8	1.6	3.7
10. Satisfied with decision approach	5-Very Satisfied	1-Very Dissatisfied	3.9	4.0	2.6	4.0
11. Satisfied with the decision you made	5-Very Satisfied	1-Very Dissatisfied	3.6	4.4	2.1	3.5
12. Comfortable applying approach to other chemical AA decisions	5-Strongly Agree	1-Strongly Disagree	3.2	3.6	2.4	3.7
13. Likely that your institution would use the approach	5-Very likely	1-Very Unlikely	3.5	3.4	2.9	3.3

*The numbers indicate the average response to the question listed, based on the specified 5 point scales. Eleven out of 12 participants completed the “unaided” survey; 5 out of 5 completed the “individual MAVT” survey; 7 out of 7 completed the “group MAVT” survey; and 11 out of 11 completed the “SDM” survey.

DOI: 10.1002/ieam.4316

Table 4. Session 2 alternative selections

Participant	Individual ranking	Final selection
Group 1 (individual exercise)		
1	Armor99	Armor99
2	Aquaslide	Armor99
3	HullSaver	Armor99
4	Aquaslide	Armor99
5	Armor99	Armor99
Group 2 (group exercise)		
6	Barrier	Barrier
7	Aquaslide	Barrier
8	Aquaslide	Barrier
9	HullSaver	Barrier
10	Barrier	Barrier
11	StreamXL	Barrier
12	Barrier	Barrier

were able to explore the effect of changing weights or value functions. In the initial ranking exercise the participants preferred several different alternatives (see Table 4). Following individual MAVT analysis and sensitivity analysis and deliberation, each participant independently selected Armor99 as their preferred alternative. As Table 3 shows, in terms of the decision process, the individual MAVT was rated higher than the preworkshop unaided exercise on many dimensions.

The individuals in Cohort 2, the lightly-facilitated group MAVT treatment, fared differently. The group used the same Decerns MCDA software; however, the group rejected the highest scoring alternative, Armor99. During group deliberation, participants expressed concern that because individuals' weights were aggregated, they may not accurately reflect the group's values. Individuals also disliked the compensatory nature of MAVT, which allows poor performance on 1 criterion to be offset by superior performance on another. The participants in Cohort 1 (the individual MAVT treatment) may also have disliked the compensatory nature, but they had more leeway to explore the effect of changing their value functions in the software, something the individuals in Cohort 2 were not able to do. The group instead used a simple aggregated ranking method to reach group consensus for "Barrier." Each alternative received points for each first (4 points), second (3 points), or third (2 points) place ranking it received in the individual ranking exercise at the start of the session. The points were totaled for each alternative to generate the total score, and the highest scoring alternative was selected.

Table 3 shows that the participants evaluated the individual and group MAVT exercises quite differently, rating the group MAVT approach much lower than the unaided approach on most criteria. Despite rejecting the MAVT analysis' outcome, group MAVT participants did report that the process increased understanding of the information, trade-offs, and their values.

Session 3: Group decision making aided by SDM. The participants then regrouped for a more intensively facilitated plenary session with SDM. Participants reported appreciating the opportunity to explore trade-offs between alternatives in a group setting before applying weights. Several noted that taking the time to understand the relative performance of each alternative helped them to better understand trade-offs. Participants rated SDM higher than unaided decision making on improving understanding of information, trade-offs, and their values. Others liked how group members were able to express their weights individually and then "drill down" into the calculations as a group to explore how the range of individual weights impacted overall MAVT value scores. This allowed the group to use the MAVT results to debate and come to a consensus on a preferred option, rather than relying on 1 set of weights as in the group MAVT exercise. However, several participants noted that SDM would be challenging to implement in their organization since it requires a skilled facilitator and more time from participants. Additionally, participants indicated lower satisfaction with the group SDM decision compared to the individual MAVT session (3.5 vs 4.4), despite equal ratings for the 2 decision approaches (4.0). Three of the 4 participants who ranked the SDM decision lower than individual MAVT did so based on dissatisfaction with methodological features (weighting and use of linear value functions). The fourth participant was dissatisfied with the selected alternative. Aside from these points, the participants' ratings for the SDM exercise were mostly similar to the individual MAVT exercise.

DISCUSSION

Diversity of "unaided" decision methods but convergence in decisions

Current AA guidelines lack specific guidance on how to make decisions, so it is useful to note the wide variety in decision methods the participants used for the individual preworkshop unaided AA exercise. Even though all participants had considerable experience in decision making concerning product alternatives, their approaches varied widely. Despite this variety, a large majority arrived at the same choice: 8 of 12 participants chose Zn2000, and a 9th chose Zn2000 as 1 of 2 top choices. Only 2 participants chose other options, and 1 participant did not report his choice of the 9 who chose Zn2000, 4 used a narrative approach, 2 used a rule-based approach, 2 used a combined rule-based and narrative approach, and 1 used an MCDA-like method. The 2 participants who did not choose Zn2000 used a rule-based or a rule-based and narrative approach. Given the diversity of decision-making approaches among Zn2000 adopters (and the use of rule-based and rule-based narrative methods by both Zn2000 and nonZN2000 adopters), it is unlikely that the decision-making approach itself played a substantial role in biocide choice. Review of the participants' descriptions of their respective decision processes indicates that their weighting of the decision criteria played an important role. The 9 who emphasized technical performance and cost tended toward Zn2000; the 2 who placed equal or greater emphasis on human health and environmental concerns selected other biocides.

MAVT and SDM substantially enhance understanding of the problem

Understanding and resolving trade-offs among alternatives is essential in AA. As the number of criteria grows, it becomes increasingly difficult to evaluate trade-offs. MAVT and SDM methods are often proffered as effective tools to support the evaluation of trade-offs (Kiker et al. 2005). The results are consistent with this view. Participants reported substantially higher understanding of the trade-offs between alternatives after the individual MAVT (4.4) and SDM (4.0) exercises, relative to their unaided decision approach (3.3). Some participants found the visualizations of the trade-offs most helpful:

- "Being able to see the trend lines…really helped me see how different products fare." Participant 9 (MAVT—Individuals)
- "DECERNS allowed us to visually see the effect of each criterion (*on the performance of alternatives)*." Participant 12 (MAVT—Individuals)
- "I like the ability to be able to see the impact of weighting the different criteria on the alternative ranking (e.g., human health, ecotox, performance, cost). It's a great comparative tool for visualizing trade-offs." Participant 3 (SDM)
- "It helped me to visualize the trade-offs in a way that balanced inherent importance with potential to change." Participant 5 (SDM).

In actual AA situations, it is often important for decision makers to communicate and justify the chosen alternative to various stakeholders. One benefit often claimed for formal methods such as MAVT or SDM is that they can facilitate communication about the choices made. Our results modestly support this belief. Participants felt that the individual MAVT and the SDM approaches would enable or promote transparency more than their unaided approach would (both 3.8 vs 3.4 for unaided) and would help them communicate the results and decision rationale more (both 4.0 vs 3.7 for unaided). On the other hand, it is perhaps surprising that the effects are not stronger.

Group vs individual decision making

It is unlikely that an individual will perform AA alone; the AA process requires input from a variety of disciplines. Likewise, stakeholders affected by the ultimate decision will no doubt engage with the AA process as well. However, the trade-off analysis at the core of AA—deliberation over and selection of the preferred alternative—could be done either by an individual or by a group. This raises the question of the relative usefulness of various decision analysis methods in group and individual settings. The evidence from our workshop is mixed in that regard.

Compare the results between the individual and group MAVT sessions. Recall that these 2 exercises were done in parallel, based on the same case study and using the same software. Four out of the 5 participants in the individual MAVT exercise chose the top performing alternative found using Decerns MCDA as their preferred alternative. The group MAVT participants entirely rejected the alternative suggested as superior by the software. On some metrics, such as improving understanding of available information and of the participants' own values, individual and group MAVT still scored similarly. On other metrics, the 2 approaches fared very differently. The participants in the group MAVT did not feel it helped with transparency or communication, the outcome did not feel right to them, they were not satisfied with the approach or the decision, and would not be comfortable applying this to other AA decisions.

We cannot definitively attribute this difference between the individual and group MAVT to any one factor, as this was far from a randomized controlled trial. We can speculate about several factors that may have contributed. First, the participants in the individual treatment enjoyed the fact that they could interact directly with the software and explore what happened if they changed weights and other inputs (sensitivity analysis). Among the group exercise participants, some reported that the tool felt more like a black box since they were unable to individually adjust weights and settings to explore outcomes. Unlike the individual exercise group, they did not have time to personally familiarize themselves with the software. A more extensively facilitated version, where the group collectively explored the effects of different weights, might have been better-received. Another alternative is a hybrid approach where participants spend some time exploring individually and then discuss their findings in a more extensively facilitated group setting. Second, the participants used their own weights in the individual MAVT exercise. In the group setting, the facilitator had to average the weights across individuals since Decerns MCDA, and MAVT in general, are not designed to consider multiple sets of weights. As a result, participants did not necessarily feel that the group weights reflected their own weights. A more flexible and perhaps more heavily facilitated version of this group approach might have avoided this pitfall. Third, some participants in the group MAVT were not comfortable with the decision framework used, i.e., a simultaneous framework in which all criteria were considered at once and in which good performance on one criterion could compensate for poor performance on another. The individual MAVT participants also used a simultaneous framework to start, but they could individually restructure it such that certain criteria were used to screen out some alternatives before applying the simultaneous framework to the rest. Fourth, in the group discussion, 1 participant was strongly critical of the specific version of MAVT used, arguing that additive weighting was fundamentally incompatible with the analytical needs of the problem and that linearized value functions distorted the evidence. Their articulate and vigorous arguments may have contributed to the group rejecting the outcome.

From this, one might tentatively prefer the individual over the group approach. However, comparing the individual MAVT with the SDM results does not support such a conclusion. The SDM exercise was also conducted in a group, but with more extensive facilitation than the MAVT group exercise. Individuals' weights were preserved throughout the SDM process, and there was far more structured discussion about those weights in the SDM exercise than the relatively simple averaging that occurred in the MAVT group setting. The SDM exercise also involved more discussion around alternatives prior to the weighting exercise, eliminating options that the group felt were dominated by others, and constantly seeking consensus before moving to a next step. The scores for the SDM exercise were generally close to those of the individual MAVT, and the participants were equally satisfied with both approaches (both 4.0). Adopting an appropriate facilitation approach and managing group dynamics emerged as important factors in developing a successful MCDA-based approach to an AA decision in the group setting.

The participants were less satisfied with the decision resulting from the SDM exercise (3.5) than from the individual MAVT (4.4). This may reflect dissatisfaction with the weighting method rather than with SDM more generally. While participants rated the difficulty of SMARTER weighting (used with the SDM case) to be similar to swing weighting (used with MAVT), more than one-half of the responding participants specifically expressed discontent with the SMARTER weighting process. For example:

- "I was least satisfied with the potential for the SMARTER weighting to heavily weight an alternative with only 1 metric (e.g., cost)." Participant 5 (SDM)
- "The way the weighting was done was somewhat opaque." Participant 6 (SDM).

Additionally, some participants wanted a second round of weighting to allow them to revise their weights given what they learned from the discussion following the weighting exercise. They felt that the MAVT-based value scores with which the group made a final decision reflected their initial weights, not their revised thinking. While the opportunity to revise weights based on new information is typically part of an SDM process, this step was omitted due to time constraints. Choosing an appropriate method to elicit weights, providing participants an opportunity to discuss their rationales for their weights, allowing participants to experiment with the effects of varying those weights (i.e., sensitivity analysis), and allowing time for multiple rounds of discourse and weighting appear to be important factors regarding satisfaction with the decision process. If an AA decision process is conducted in a group setting, and if it is considered important for the group to support the eventual decision, then the facilitation approach should be sufficiently flexible to allow for the wide range of unpredictable events that can occur in group settings. If some members of a group feel that a decision is being imposed upon them by the method used, they are less likely to support the final outcome. A facilitation approach that surfaces potential resistance from individuals and that includes enough time to explore the underlying causes of individuals' resistance and to adapt the decision approach in response is more likely to be accepted by the group. Such flexibility is of course not a guarantee of success, but it is plausible that a more extensively facilitated version of the group MAVT exercise, with the capacity to adjust weighting schemes and other aspects of the decision process, would have been better-received.

Satisfaction, comfort, likelihood of use, alignment with expectations, and defining "success"

Setting aside the group MAVT exercise, the participants were roughly equally satisfied with their unaided method, the individual MAVT, and the SDM method (3.9, 4.0, and 4.0, respectively). Similarly, the participants reported similar scores for the extent that the top-performing alternative reflected what mattered to them in the unaided exercise (3.9), the individual MAVT (3.8), and the SDM (3.7). The participants did appear somewhat more comfortable applying individual MAVT (3.6) and SDM (3.7) to other chemical AA decisions than their unaided method (3.2), but that did not translate into higher likelihood of use: participants predicted that their respective unaided methods were marginally more likely (3.5) than individual MAVT (3.4) or SDM (3.3) to be adopted by their respective organizations. This discrepancy does not appear to be driven by concerns that MAVT or SDM are too difficult or complex. Participants rated individual MAVT (4.0) and SDM (3.8) as slightly less difficult than their unaided methods (3.5). The main barrier to adoption cited for individual MAVT was the affordability and availability of software. For SDM, participants identified the time and cost of training and facilitation as the leading barriers to adoption.

Some of these discrepancies raise questions about how to define "success" in applying MCDA to AA decisions. In assessing the quality of decision-making processes, the decision science literature typically relies on many of the metrics used in this study, including enhanced understanding of information, trade-offs, and values; level of transparency; improved capacity to explain the basis of the decision; and so on (Belton and Stewart 2002; Kiker et al. 2005). We observed individual MAVT and SDM scoring better than unaided methods on many of these metrics but only marginally better in terms of satisfaction with the overall approach. Individual MAVT and SDM even scored slightly lower than the unaided method on the extent to which the top-performing alternative reflected what is important to participants or how likely participants were to use the method. In order to determine whether a more structured approach to AA decisions is valuable, we cannot simply ask participants whether they are more satisfied with their decision or with their approach. There is also no objective "best" outcome in most AA decisions to use as a benchmark. Although our survey results suggest that MAVT and SDM add value in a number of dimensions, we cannot point to 1 approach as being "better" than the other. Therefore, AA practitioners might need to explore different weighting methods and decision approaches in any given situation in order to understand how robust the respective outcomes are, as consistent with research on MCDA application in other contexts (Salminen et al. 1998; Yatsalo et al. 2007).

The findings of the workshop indicate that decision making, the final element of the AA process, will be challenging for stakeholders engaged in programs such as those in California or the European Union. Such programs typically focus on the early stages of the AA process (identifying possible alternatives and criteria of concern, as well as data gathering) but provide less guidance on tools for reaching decisions. Yet the decisions are likely to be complex, considering a large number of variables and alternatives. Participants in this workshop reported that tools such as MCDA could be helpful in organizing and presenting information, despite the limited nature of the case study used. Nonetheless, participants felt that these tools would be unlikely to be used in their organizations due to issues such as cost and lack of familiarity. The current work makes clear that much should be done to increase the level of comfort and accessibility of decision-making tools. That could include developing more extensive guidance on the use of MCDA in AA settings, development of easy-to use AA decision support tools, and training on their use. Published case studies on AA would be valuable, and they should include a discussion of how the alternatives were evaluated. Even if the process used was essentially narrative, a formal description of the decision process can help others identify which MCDA approach is most suitable for their context. There is also a role for professional societies, including the Society for Risk Analysis and the newly formed Association for the Advancement of Alternatives Assessment.

Our study has several limitations. The sample size was small due to practical and funding limitations so it is difficult to generalize the results. We included only a few AA decision-making tools due to time constraints; a larger and broader survey of tools would be worthwhile. All of our participants were college educated, had technical backgrounds in sustainability and product stewardship, and had at least some familiarity with AA. This may not be the case for all potential stakeholders involved in AA decision making. Stakeholders who are less familiar with concepts such as MAVT and SDM might display greater reluctance to use such tools. On the other hand, our participants came from different viewpoints that would likely be representative of individuals who would lead AA efforts and make decisions about scoping, tools, training, etc.

CONCLUSIONS

Our primary objective was to explore whether MCDA methods would help decision makers in AA improve their understanding of the problem, including their values and the trade-offs, as MCDA is known to do in other contexts. Our results clearly suggest that decision makers did report greater understanding of their values, the information provided, and the trade-offs involved, after using an MCDA approach relative to their unaided decision. However, the different MCDA approaches, and the group vs individual settings, gave rise to different experiences with respect to other attributes such as transparency and likelihood of adoption. Below we return to the 3 stylized sample AA scenarios we sketched in the introduction and use our results to draw tentative recommendations for each decision context.

In Scenario 1, an individual is performing an AA for their company's own internal use. In this context, tools that support group dialogue, transparency, and the ability to communicate the logic behind a decision to external audiences is not critical. Individuals charged with conducting an AA within a large organization may prefer a method that is simple, yet allows them to articulate weights, investigate performance and trade-offs, and explore sensitivity. Our individual MAVT treatment demonstrates that participants valued the ability to directly interact with and perform sensitivity analyses with off-the-shelf MCDA software, leading to greater satisfaction with the ultimate decision. In the context of an individual performing an AA within a large company, off-the-shelf MCDA software (such as Decerns MCDA) may be preferred over formal group decision-making approaches (i.e., group MCDA or SDM).

In Scenario 2, a trade association is conducting the AA. Transparency and being able to communicate the results are more important in this scenario, while the satisfaction of each individual with the decision is less critical. The SDM treatment in our workshop may be better suited for this scenario, particularly given its ability to incorporate values from multiple stakeholders into the analysis.

In Scenario 3, a government agency is conducting the AA. The outcome will likely be contested in various venues, so the process must be legally defensible, inclusive, and transparent. Individual satisfaction with the decision is not an objective in this scenario, but transparency and the inclusion of stakeholder perspectives and values becomes critical. In this setting, multiple approaches could be considered, depending on the desired level of engagement with stakeholders, and on more practical considerations, such as physical proximity of the relevant stakeholders. An agency may opt to separate the consultation process from the analysis. In this scenario, an individual would conduct an MAVT, but they would select appropriate criteria, alternatives, and weights based on input received through stakeholder consultation. Alternatively, multiple individuals (within the agency, or across the agency and stakeholder groups) may conduct their own MAVT analysis, with the agency synthesizing across analyses to come to a decision.

 DOI: 10.1002/ieam.4316

Finally, an agency may decide to conduct a group based SDM process with agency and stakeholder experts to explore and select from a set of alternatives. This work provides insight into the multiple ways in which MCDA methods can help decision makers in various AA settings, as well as insight into which MCDA methods may be most appropriate given the decision context and the objectives of the organization leading the analysis. These scenarios are meant to illustrate which MCDA approaches may be preferred in various decision contexts based on our findings. Since it is not possible to fully capture the details and variations of all contexts in which AAs may be applied, and given our small sample size and limited number of methods and treatments tested, these are meant as guidance rather than firm recommendations on best practices.

Acknowledgment—C Beaudrie and T Malloy received funding to support this project from the Society for Risk Analysis New Initiatives grant program. T Malloy received funding to support this project from the UC Center for Environmental Implications of Nanotechnology, which is funded by a cooperative agreement from the National Science Foundation and the EPA (NSF DBI-0830117; NSF DBI-1266377). The authors thank the workshop participants, and Northwest Green Chemistry for its assistance in providing data and background information.

Disclaimer—The authors declare no conflicts of interest.

Data Availability Statement—Data and associated metadata and calculation tools are available upon request by contacting corresponding author Timothy Malloy (Malloy@law.ucla.edu).

SUPPLEMENTAL DATA

Case study narrative.

Figure S1. Participant survey results.

Table S1. Comparison of decision analysis methods

Table S2. Participant unaided decision methods

ORCID

Christian Beaudrie http://orcid.org/0000-0003-3272-0371
Charles J Corbett http://orcid.org/0000-0003-1814-3977
Thomas A Lewandowski http://orcid.org/0000-0003-1717-4296
Timothy Malloy http://orcid.org/0000-0002-6521-4005
Xiaoying Zhou http://orcid.org/0000-0002-7901-2842

REFERENCES

Arvai J, Gregory R. 2003. Testing alternative decision approaches for identifying cleanup priorities at contaminated sites. *Environ Sci Technol* 37:1469–1476.

Arvai JL, Gregory R, MacDaniels TL. 2001. Testing a structured decision approach: Value-focused thinking for deliberative risk communication. *Risk Anal* 21:1065–1076.

Beaudrie C, Schroeder C. 2017. Swing-weighting tool [web application]. Compass Resource Management Ltd. [accessed 2017 Dec]. http://www.compassrm.com/decisiontools

Bell ML, Hobbs BF, Elliott EM, Ellis H, Robinson Z. 2001. An evaluation of multicriteria methods in integrated assessment of climate policy. *J Multi-Crit Decis Anal* 10:229–256.

Bell ML, Hobbs BF, Ellis H. 2003. The use of multicriteria decision-making methods in the integrated assessment of climate change: Implications for IA practitioners. *Socio-Econ Plan Sc* 37:289–316.

Belton V, Stewart TJ. 2002. Multiple criteria decision analysis: an integrated approach. 1st ed. Boston (MA): Kluwer Academic. 161 p.

Bessette DL, Arvai J, Campbell-Arvai V. 2014. Decision support framework for developing regional energy strategies. *Environ Sci Technol* 48: 1401–1408.

Booker J. 2011. A survey-based methodology for prioritizing the industrial implementation qualities of design tools. *J Eng Design* 7:507–525.

Cegan JC, Filion AM, Keisler JM, Linkov I. 2017. Trends and applications of multicriteria decision analysis in environmental sciences: Literature review. *Environ Syst Decis* 37:123–133.

Cinelli M, Coles SR, Kirwan K. 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. *Ecol Indic* 46:138–148.

[DTSC] Department of Toxic Substances Control. 2013. California code of regulations, title 22, division 4.5. Safer consumer products. [accessed 2020 Jan 20]. https://dtsc.ca.gov/wp-content/uploads/sites/31/2018/07/SCP-Final-Regs-Text-10-01-2013.pdf

[DTSC] Department of Toxic Substances Control. 2017. Alternatives analysis guide version 1.0: (US). Sacramento (CA): DTSC. 235 p. https://dtsc.ca.gov/wp-content/uploads/sites/31/2018/04/DTSC-Alternatives-Analysis-Guide-Version-1-0-June-2017.pdf

Eason T, Meyer D, Curran M, Upadhyayula V. 2011. Guidance to facilitate decisions for sustainable nanotechnology. Cincinnati (OH): Environmental Protection Agency. [accessed 2018 Jun 19]. https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NRMRL&dirEntryId=238589

[ECHA] European Chemicals Agency. 2010. General approach for prioritisation of substances of very high concern (SVHCS) for inclusion in the list of substances subject to authorization. Helsinki (FI). 13 p. https://echa.europa.eu/documents/10162/13640/gen_approach_svhc_prior_in_recommendations_en.pdf

Edwards W, Barron F. 1994. SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement. *Organ Behav Hum Decis Process* 60:306–325.

[EPA] Environmental Protection Agency. 2018. Remedial investigation/feasibility study and selection of remedy. 40 C.F.R., Sect. 300.430. Washington (DC): Government Printing Office.

French S. 1989. Readings in decision analysis. Boca Raton (FL): Chapman & Hall/CRC. 224 p.

Gregory R, Failing L, Harstone M, Long G, McDaniels T, Ohlson D. 2012. Structured decision making: A practical guide to environmental management choices. Oxford (UK): Wiley-Blackwell. 315 p.

Gregory R, Wellman K. 2001. Bringing stakeholder values into environmental policy choices: A community-based estuary case study. *Ecol Econ* 39:37–52.

Guitouni A, Martel J-M. 1998. Tentative guidelines to help choosing an appropriate MCDA method. *Eur J Oper Res* 109:501–521.

Hajkowicz S. 2007. A comparison of multiple criteria analysis and unaided approaches to environmental decision making. *Environ Sci Pol* 10:177–184.

Huang IB, Keisler J, Linkov I. 2011. Multicriteria decision analysis in environmental sciences: Ten years of applications and trends. *Sci Total Environ* 409:3578–3594.

Jacobs MM, Malloy TF, Tickner JA, Edwards S. 2016. Alternatives assessment frameworks: Research needs for the informed substitution of hazardous chemicals. *Environ Health Perspect* 124:265–280.

Kain J-H, Söderberg H. 2008. Management of complex knowledge in planning for sustainable development: The use of multicriteria decision aids. *Environ Impact Assess Rev* 28:7–21.

Keeney RL, Raiffa H. 1972. A critique of formal analysis in public sector decision making. In: Drake AW, Keeney RL, Morse PM, editors. Analysis of public systems. Cambridge (MA): MIT Press. p 64–74.

[KemI] Swedish Chemicals Agency. 2007. The substitution principle. Sundbyberg (SE): KemI. Report Nr 8/07. [accessed 2015 Mar 15]. https://www.kemi.se/global/rapporter/2007/rapport-8-07-the-substitution-principle.pdf

Kiker GA, Bridge TS, Varghese A, Seager TP, Linkov I. 2005. Application of multicriteria decision analysis in environmental decision making. *Integr Environ Assess Manag* 1:95–108.

Lahdelma R, Salminen P, Hokkanen J. 2000. Using multicriteria methods in environmental planning and management. *Environ Manag* 26(6):595–605.

Lindahl M. 2006. Engineering designers' experience of design for environment methods and tools—Requirement definitions from an interview study. *J Clean Prod* 14(5):487–496.

Linkov I, Moberg E. 2012. Multicriteria analysis: Environmental applications and case studies. Boca Raton (FL): CRC Press. 204 p.

Linkov I, Varghese A, Jamil S, Seager TP, Kiker GA, Bridges TS. 2004. Multicriteria decision analysis: A framework for structuring remedial decisions at contaminated sites. Dordrecht (NL): Kluwer Academic. 15 p.

Malloy T, Blake A, Linkov I, Sinsheimer P. 2015. Decisions, science, and values: Crafting regulatory alternatives analysis. *Risk Anal* 35:2137–2151.

Malloy TF, Sinsheimer PJ, Blake A, Linkov I. 2013. Use of multicriteria decision analysis in regulatory alternatives analysis: A case study of lead free solder. *Integr Environ Assess Manag* 9:652–664.

Malloy TF, Zaunbrecher VM, Batteate CM, Blake A, Carroll WF, Corbett CJ, Hansen SF, Lempert RJ, Linkov I, McFadden R et al. 2017. Advancing alternative analysis: Integration of decision science. *Environ Health Perspect* 125:066001.

National Academy of the Sciences, National Research Council. 2014. A framework to guide selection of chemical alternatives. Washington (DC): National Academies Press. 280 p.

Németh B, Molnár A, Bozóki S, Wijaya K, Inotai A, Campbell JD, Kaló Z. 2019. Comparison of weighting methods used in multicriteria decision analysis frameworks in healthcare with focus on low-and middle-income countries. *J Comp Effect Res* 8(4):195–204.

Northwest Green Chemistry. 2017. Washington state antifouling boat paint alternatives assessment report. Seattle (WA). 176 p.

Ogunseitan OA. 2016. Power failure: The battered legacy of leaded batteries. *Environ Sci Technol* 50:8401–8402.

Rossi M, Peele C, Thorpe B. 2012. BizNGO chemical alternatives assessment protocol. Somerville (MA): BizNGO. [accessed 2017 Dec 11]. https://www.bizngo.org/resources/entry/chemical-alternatives-assessment-protocol

Salminen P, Hokkanen J, Lahdelma R. 1998. Comparing multicriteria methods in the context of environmental problems. *Eur J Oper Res* 104:485–496.

Steele K, Carmel Y, Cross J, Wilcox C. 2009. Uses and misuses of multicriteria decision analysis (MCDA) in environmental decision making. *Risk Anal* 29:26–33.

Tickner JA, Schifano JN, Blake A, Rudisill C, Mulvihill MJ. 2015. Advancing safer alternatives through functional substitution. *Environ Sci Technol* 49:742–749.

Tsang MP, Bates ME, Madison M, Linkov I. 2014. Benefits and risks of emerging technologies: Integrating life cycle assessment and decision analysis to assess lumber treatment alternatives. *Environ Sci Technol* 48:11543–11550.

Wilson C, McDaniels T. 2007. Structured decision-making to link climate change and sustainable development. *Clim Policy* 7:353–370.

Wilson RS, Arvai JL. 2006. Evaluating the quality of structured environmental management decisions. *Environ Sci Technol* 40:4831–4837.

Yatsalo B, Gritsyuk S, Sullivan T, Trump B, Linkov I. 2016. Multicriteria risk management with the use of DecernsMCDA: Methods and case studies. *Environ Sys Decis* 36:266–276.

Yatsalo BI, Kiker GA, Kim J, Bridges TS, Seager TP, Gardner K, Satterstrom FK, Linkov I. 2007. Application of multicriteria decision analysis tools to two contaminated sediment case studies. *Integr Environ Assess Manag* 3:223.

Zheng Z, Peters GM, Arp HPH, Andersson PL. 2019. Combining in silico tools with multicriteria analysis for alternatives assessment of hazardous chemicals: A case study of decabromodiphenyl ether alternatives. *Environ Sci Technol* 53:6341–6351.

Zhou X, Schoenung JM. 2008. An integrated impact assessment and weighting methodology: Evaluation of the environmental consequences of lead-free solder alternatives. In: Proceedings of IEEE Electron Environment; 2008 May 18; San Francisco, CA.

