Navigating inconsistent preferences: A multimethod approach to making informed decisions

David M. Martin

Maryland/DC Chapter, The Nature Conservancy, Bethesda, Maryland

Correspondence
David M. Martin, 425 Barlow Place, Suite 100, Bethesda, MD 20814.
Email: david.martin@tnc.org

Abstract
This study presents a decision-analytic framework for prioritizing conservation strategies. The framework is based on combining direct and indirect preference-elicitation methods and analyzing inconsistencies between the methods. A case study with The Nature Conservancy's Chesapeake Bay (The United States) agriculture team is presented. Participants evaluated six strategies to engage with agribusinesses, farmers, and farm landowners and increase adoption of nutrient and soil conservation and stream and wetland restoration activities. Impact, feasibility, and risk criteria and performance measures were developed to compare the strategies. Participants individually evaluated the strategies using a multimethod approach. One method included direct ranking based on an intuitive assessment of the strategies. The second method included indirect ranking based on swing weighting and multicriteria analysis. Some participants made adjustments to reduce inconsistencies between the methods. Results show that final rankings were more consistent than initial rankings. Inconsistencies can be reduced by understanding potential bias in the direct method and clarifying assumptions in the indirect method. This study provides evidence that a multimethod approach can deliver useful insights to inform decisions.

KEYWORDS
Chesapeake Bay, conservation, decision analysis, elicitation, multicriteria analysis

1 | INTRODUCTION

Environmental managers often must develop a methodology to assess and evaluate hard decisions. Such a methodology may include formal or informal approaches that obtain information from knowledgeable participants. Any approach may incite challenges for decision makers as well as participants. For instance, the governance structure between decision makers and participants is often a sensitive matter (Straus, 2002). Different values or concerns may be held by all persons involved in the decision (Keeney, 1992), and there are challenges in obtaining and deliberating those values (Bond, Carlson, & Keeney, 2008). There are often conflicting preferences on criteria that characterize values (Slovic, 1995) and difficulties in separating values from facts (Von Winterfeldt, 2013). Developing some understanding of the potential consequences of decisions matter, but predictive models provide only a partial understanding of decision outcomes due to various types of uncertainty.
(Fischhoff & Davis, 2014). There are difficulties in communicating preferences because human behavior and cognition cause mental shortcuts and bias (Kahneman, 2013). Further, decisions are often made in the absence of a structured elicitation approach that allow participants to think carefully about options and provide meaningful value-based judgments to decision makers (Gregory et al., 2012).

In collaborative decision contexts, participants are asked to give their preferences on the desirability of management alternatives (McDaniels, 2019). Decision analysis is a systematic and transparent approach to facilitate the preference-elicitation process and pinpoint areas of agreement and disagreement among participants (Keeney, 2004). Decision analysis includes a set of five principles: (a) develop an adequate description of the decision; (b) develop a common understanding of the fundamental values or concerns from decision makers or other knowledgeable participants; (c) develop distinct alternative courses of action to choose from; (d) evaluate the alternatives on their characteristics, which are directly representative of the fundamental values; and (e) evaluate tradeoffs between the alternatives. The first three are sometimes referred to as front-end development principles, whereas the latter two are referred to as back-end evaluation principles (Keeney, 2020).

This article aims to inform the back-end evaluation principles of decision analysis as they are used to aid decision makers in gathering input from participants prior to making decisions. Different evaluation methods are used in decision analysis. A direct method allows participants to intuitively evaluate the alternatives and their values. It is assumed that their evaluation represents tradeoffs between the alternatives, but it is very hard to observe this behavior because of personal biases and irrational thinking (Bessette, Wilson, & Arvai, 2019; Slovic, Fischhoff, & Lichtenstein, 1976). An indirect method attempts to break an intuitive process down into parts that can be analyzed and integrated into a comparable, and often quantitative, evaluation (Von Winterfeldt & Edwards, 1986). The intent of the indirect approach is to develop an explicit evaluation of tradeoffs between the alternatives. It is believed that focusing on tradeoffs will reduce bias while also providing useful insights into the desirability of the alternatives. One evaluation method is not better than another, and any method can yield insights for decision making. However, inconsistent preference judgments are well-documented under both methods (Bessette, Wilson, & Arvai, 2019).

This research emphasizes a multimethod approach, which can reduce inconsistencies between methods and increase the validity of results relative to using any single method. The intent of a multimethod approach is to gain insights toward a comprehensive understanding of decision-relevant preferences. There is evidence that this approach results in transparency, accountability, and a more accurate reflection of preferences in experimental (Arvai & Gregory, 2003; Hobbs & Horn, 1997; Hostmann, Bernauer, Mosler, Reichert, & Truffer, 2005) and reflective analyses (Bessette, Wilson, Beaudrie, & Schroeder, 2019; Gregory et al., 2012; Keeney, Von Winterfeldt, & Eppel, 1990). There is a strong need for real-world case studies that demonstrate broad applicability and further evidence in many environmental management contexts.

### 1.1 Case study

The Chesapeake Bay is the largest estuary in North America covering 11,601 km$^2$. In 2010, the United States Environmental Protection Agency established a 2025 water quality mandate for nutrient (nitrogen and phosphorus) and sediment management in the Chesapeake Bay watershed (U.S. Environmental Protection Agency, 2010). Agriculture, including cropland and animal operations such as poultry, dairy, and beef cattle, contributes the greatest amount of nutrient loads to the Chesapeake Bay (Chesapeake Assessment and Scenario Tool, 2020). Farmers and agricultural landowners have implemented conservation practices such as planting cover crops and matching fertilizer applications to crop needs, which combine to reduce sediment and nutrients leaving fields and animal operations, and stream and wetland restoration. These practices have reduced nitrogen loading to the Chesapeake Bay by 38% between 1985 and 2018 (Chesapeake Bay Commission, 2019). These practices need to reduce nitrogen loading to the Chesapeake Bay by 72% between 2018 and 2025 (Chesapeake Bay Commission, 2019), which will require accelerated implementation of these conservation practices. A significant increase in agricultural conservation work is critical.

The Nature Conservancy is a global non-profit organization whose mission is to conserve the lands and waters on which all life depends. The Nature Conservancy’s Chesapeake Bay agriculture team aims to reduce nutrients and sediment from agricultural landscapes at a scope and scale that restores aquatic habitat in the Chesapeake Bay and supports a thriving agricultural community. The team recently engaged in a strategic planning process to allocate staff time and resources across the watershed. Six possible strategies were developed over several months of discussion and evidence-based evaluations. The context for the strategies is based on engagement activities that are expected to encourage farmer and farm landowner
(leasing agent) adoption of conservation practices and make a meaningful contribution to the Chesapeake Bay water quality mandate. The strategies are referred to as Strategy 1, Strategy 2, Strategy 3, Strategy 4, Strategy 5, and Strategy 6 to avoid disclosing details of internal strategies publicly. Further description of the strategies and their proposed activities is beyond the scope given space constraints, as the focus of this article is on evaluating decision-analytic methods, and therefore depth of understanding about the strategies is not necessary.

The six strategies are feasible options to address the team’s goals for the project. However, all strategies cannot be implemented at one time for time and feasibility reasons. Therefore, the team’s project managers identified the need for some analysis to aid in their decision on which strategies to move forward with implementation, which strategies should not be implemented, and which strategies should be phased in over time or delayed for future implementation. A ranking evaluation was developed to inform this prioritization context: Which strategies are the best ones to invest in now to help the team achieve the outcomes they are seeking? The project managers desired an approach to help them more consistently gather, synthesize, and reflect on the entire team’s preferences. The multimethod approach to decision analysis was used to provide this insight. This research allowed 13 team members to provide a consistent type of input to the project managers and aid in their choice of high-priority strategies.

2 | MATERIALS AND METHODS

2.1 | Objectives and criteria

In decision analysis, alternative strategies are evaluated against how well they are expected to achieve fundamental values or concerns. Objectives and criteria are developed to discern among the strategies. This study drew knowledge from best practices in conservation planning (The Nature Conservancy, 2000; Foundations of Success, 2018; Conservation Measures Partnership, 2020) and developed impact, feasibility, and risk objectives and 11 measurable criteria for the objectives (Table 1).

Three types of measures are generally used to develop some expectation on the possible performance of alternative strategies: natural, proxy, and constructed (Keeney & Gregory, 2005). Natural measures directly represent the criteria (e.g., dollars for a cost criterion). Proxy measures do not directly represent the criteria but are highly correlated with them (e.g., person-days for a cost criterion). For many criteria, it is difficult to develop a natural or proxy measure. In these cases, constructed measures are

| Objective | Criterion | Consideration |
|-----------|-----------|---------------|
| Impact    | Water quality | Is the strategy designed with intent to achieve significant contributions to water quality? |
| Climate   |            | Is the strategy designed with intent to achieve significant contributions to climate resilience and mitigation? |
| Permanency|            | If successfully implemented, will the strategy secure long-lasting outcomes through changes to the current system? |
| Evidence  |            | Is there sufficient evidence to support the intended water quality and climate outcomes? |
| Feasibility | Key stakeholders | What is the anticipated level of support and engagement by key stakeholders whose involvement is necessary to implement the strategy? |
| Staffing  |            | How many full-time-equivalent employees are necessary to implement the strategy? |
| Time      |            | How much time will it take to implement the strategy? |
| Other funds |            | What contractual, marketing, or cost-share funds are needed to implement the strategy? |
| Risk      | Unique role/niche (internal) | Could the strategy move forward and achieve the intended outcomes without the team? |
|           | Unique role/niche (external) | Are other conservation groups doing the same or similar work in the watershed? |
|           | Unintended consequences | What is the perceived or real consequences of implementing the strategy on stakeholders (e.g., private benefits, inconvenience)? |

Note: more detailed information about the criteria and how they were measured is provided in the Appendix S1.
developed, such as 1–4 scores or “Poor,” “Fair,” “Good,” and “Very Good” ratings, as long as they are meaningful, descriptive, and indicate the degree to which the associated criterion is preferred based on empirical data or expert opinion (Keeney, 1992). Constructed measures are typically easier to evaluate than proxy or natural measures, but they are less directly related to the quantity of interest.

This study developed proxy and constructed performance measures. The project managers met with the author over several sessions to develop the measures. First, a unique narrative description communicated the rationale for each strategy’s performance per criterion (select performance measures are provided in Table 2; the full evaluation table is provided in the Appendix S1). Second, the performance measures were compared across strategies per criterion and assigned a “low,” “moderate,” or “high” performance value. It was assumed that there wasn’t enough information in the descriptions to distinguish the desirability between strategies of the same performance value. It was assumed that the information in the measures for strategies assigned a “high” value were better than measures assigned a “moderate” value, and that measures assigned a “moderate” value were better than measures assigned a “low” value. These values determined the desirability of the strategies per criterion. The performance measures and values of each strategy per criterion were collected in an evaluation table, sometimes referred to as a consequence table (Gregory et al., 2012; Appendix S1).

### 2.2 Direct evaluation

Thirteen team members participated in the direct and indirect evaluations. Each participant evaluated the strategies individually with facilitation support by the author, who designed and facilitated the analysis, but did not

| Objective | Criterion | Baseline (“worst”) | Swing (“best”) |
|-----------|-----------|-------------------|----------------|
| Impact    | Water quality | Relatively limited contribution to large-scale outcome; practice type and/or its ability to inform farmer and landowner decisions will produce low water quality improvement | Relatively high contribution to large-scale outcome; practice type and/or its ability to inform farmer and landowner decisions will produce high water quality improvement |
| Climate change | Water quality | Relatively limited contribution to large-scale outcome; practice type and/or the ability to inform farmer and landowners to adopt practices will produce high levels of GHG reduction and carbon storage | Relatively high contribution to large-scale outcome; practice type and/or the ability to inform farmer and landowners to adopt practices will produce low levels of GHG reduction and carbon storage |
| Permanency | Activities likely to be long-lasting due to well-known regulations, program terms, and landowner and farmer motivation | Activities unlikely to be continued or can be easily removed due to unknown motivation for long-term implementation |
| Evidence | Limited or no evidence of farmers and/or landowners responding to strategy activities | Direct evidence of farmers and/or landowners responding to strategy activities |
| Feasibility | Level of support has not been demonstrated among stakeholders | Level of support has been demonstrated among stakeholders |
| Staffing | 3–5 full-time equivalent employees | 1–3 full-time equivalent employees |
| Time | 5–10 years | Less than 5 years |
| Other funds | Greater than $5 million | Less than $500,000 |
| Risk | We do not have any unique skills or resources to implement the strategy activities | We have unique skills or resources to implement the strategy activities |
| Unique role (internal) | There are many other conservation groups and organizations implementing the same or similar types of activities | There are no other conservation groups and organizations implementing the same or similar types of activities |
| Unique role (external) | Potential private stakeholder benefits, potential to upset stakeholders with rules and constraints, potential to negatively impact other environmental outcomes | Limited private stakeholder benefits, low risk in upsetting stakeholders with rules and constraints, low risk of other negative environmental outcomes |

Note: This table was used to assign swing weights to the impact, feasibility, and risk objectives. Criteria weights were assigned by ranking and scoring the swings from “worst” to “best” criteria values per objective (see Appendix S1).
participate in either method. Discussions of the individual evaluations were performed in group meetings following the evaluation. The evaluations were designed in a Microsoft Excel spreadsheet tool (Appendix S1). The tool was delivered to team members with instructions to provide input that best expresses their preferences. The order that participants completed the evaluations was not tracked. However, since participants were engaged in reflective discussions about their evaluations and given time to change either evaluation, it is unlikely that the order of evaluation was a significant determinant of the results.

Under the direct evaluation, participants were instructed to study the evaluation table and rank the strategies in order of their overall value. Next, participants were instructed to assign 100 points to the top-ranked strategy. They assigned between 0 and 100 points to the remaining ranked strategies in accordance with their value against the top-ranked strategy. They continued down the list of ranked strategies, entering points either the same as or lower than the one before. Participants could assign zero points to strategies or the same points to multiple strategies. The array of points assigned to the strategies was normalized to sum to unity (1) per participant. The overall values per strategy (and summary statistics) were compared with the overall values estimated in the indirect evaluation per participant and as a group (average values).

It is important to note that a direct evaluation may not be relevant to any decision. For example, decisions with a large set of alternatives require screening tools to reduce the number to a set that can be more easily evaluated by participants (e.g., Martin, Mazzotta, & Bousquin, 2018).

### 2.3 Indirect evaluation

Methods for multicriteria analysis are based on different rules for analyzing the value of a set of strategies. One of the most popular methods is an additive value function (Dyer & Sarin, 1979). The three main steps of an additive value function are: (a) scaling each disparate criteria performance value into a common metric between 0 and 1 that maintains their differentiation among strategies; (b) weighting each criterion to reflect tradeoffs among strategies; and (c) performing a weighted-average calculation using the scaled criteria values and the weights. There are many possible scaling techniques, each with different advantages and disadvantages (Martin & Mazzotta, 2018). Similarly, there are many possible weighting techniques, each with different methods for eliciting ordinal-scale, interval-scale, or ratio-scale preferences (Danielson & Ekenberg, 2016; Von Winterfeldt & Edwards, 1986).

The project managers assumed linear performance values $z_{ij}$ per criterion $j$, strategy $i$. The “low,” “moderate,” and “high” performance values were respectively assigned 1, 3, and 5 numbers. The numbers were scaled to 0–1 values using a global scaling function $x_{ij}$:

$$x_{ij} = \frac{z_{ij} - z_{ij}^*}{z_{ij}^{**} - z_{ij}^*}$$  \hspace{1cm} (1)

for all criteria $j = 1, \ldots, 11$, strategies $i = 1, \ldots, 6$; where $z_{ij}^*$ and $z_{ij}^{**}$ are the “worst” and “best” measurements for each criterion, respectively, as assigned.

The additive value function $V_i$ aggregates the scaled performance values into an overall value per strategy $i$:

$$V_i = \sum_{j=1}^{11} w_j x_{ij}$$  \hspace{1cm} (2)

for all criteria $j = 1, \ldots, 11$, strategies $i = 1, \ldots, 6$; where $w_j$ are importance weights that sum to 1.

Importance weights are scaling factors that evaluate contributions of the criteria to the desirability of the strategies (Dyer & Sarin, 1979). In general, participants assign weights based on combining the tradeoff information contained in the evaluation table with a subjective assessment of the inherent importance of the objectives and criteria. Swing weights were elicited from participants to ensure the tradeoff property of weights was adequately considered (Von Winterfeldt & Edwards, 1986).

Swing weighting consisted of a thought experiment, with additional preference-elicitation questions in the Microsoft Excel spreadsheet tool (Gregory et al., 2012; Appendix S1). Participants were provided with the range of anticipated impact, feasibility, and risk, showing the swing from “worst” to “best” narrative descriptions for all intra-objective criteria (Table 2) and inter-objective criteria. They were then asked to compare the swings and rank them in order of their importance. They assigned 100 points to the top-ranked swing. They assigned between 0 and 100 points to the remaining ranked swings in accordance with their relative importance against the top-ranked swing. They continued down the list of ranked swings, entering points either the same as or lower than the one before. Participants could assign zero points to swings or the same points to multiple swings. The array of points assigned to the swings was normalized into weights that sum to unity (1) per participant. Cumulative weights were calculated as the product of all inter-objective criteria weights and their respective objective weights.
Each participant’s cumulative weights were combined with the scaled performance values using the additive value function (Equation (2)). The overall values per strategy (and summary statistics) were compared with the overall values developed in the direct evaluation per participant and as a group (average values).

2.4 | Deliberation about inconsistencies

A summary of the initial direct and indirect overall values per strategy was shared with the participants in one-on-one meetings with the author. We discussed possible inconsistencies in the rank order of the strategies between methods by considering preferences in their direct evaluation, their cumulative weights, and tradeoffs between top-ranked strategies based on the information contained in the evaluation table. Participants were allowed to reflect on their evaluations openly and change any preferences given in their direct or indirect evaluation. One participant made changes to their direct evaluation (direct ranking and scoring) after viewing their initial results, three participants made changes to their indirect evaluation (swing weights) after viewing their initial results, and three participants made changes to their direct and indirect evaluation after viewing their initial results. Six participants did not change their direct or indirect evaluation after viewing their initial results.

3 | RESULTS

Final cumulative weights were plotted to analyze each participant’s preferences across the criteria as they apply to indirect strategy rankings (Figure 1). Weight distributions show that the impact criteria were weighted higher than the feasibility and risk criteria, which was consistent with the initial weights (Figure S1). Final direct and indirect strategy rankings were plotted in a diagram to analyze each participant’s consistency between methods (Figure 2). In general, Strategies 1, 5, and 6 received higher value and Strategies 2, 3, and 4 received lower value across direct and indirect methods. The range and variance of direct values were greater than indirect values per strategy, and this was observed when comparing the strategy ranks between methods as well, with the exception of Strategy 1 (Appendix S1). The variance of indirect values and ranks was particularly low for Strategies 2, 3, and 6 as compared with the direct evaluation, which corresponds to agreement on these strategies across participants. Less disagreement in the indirect method is consistent with previous studies (Gardiner & Edwards, 1975; Hobbs & Horn, 1997).

Final rankings (Figure 2) were more consistent than initial rankings (Figure S2). Participants who made changes to their direct evaluation improved their consistency by two orders of magnitude, as defined by the distance between the direct and indirect ranks per strategy (e.g., Participant 13; Figure 3, upper panels). Participants who made changes to their indirect evaluation improved their consistency between two and four orders of magnitude (e.g., Participant 12; Figure 3, lower panels). Participants who made changes to both direct and indirect evaluations improved their consistencies between zero and four orders of magnitude (Figure 2; Figure S2).

4 | DISCUSSION

The main finding from this study is that navigating inconsistencies in preferences between direct and indirect methods improved participant and decision-maker insights. Participants who did not make changes to their direct or indirect evaluation were either consistent in their preferences and very insightful from the start, or inconsistent in their preferences because of strong beliefs. Strong beliefs came in the form of bias and incomplete criteria. For instance, Participant 11 was inconsistent in their preferences by 14 orders of magnitude (Figure 2). Bias for some of the strategies (e.g., emotional response to specific strategy activities) and criteria that were not included in the evaluation table (e.g., funding the strategies during the COVID-19 pandemic) were the main implications of this inconsistency. This participant agreed with the swing weighting and affirmed that the sources of inconsistencies were in their direct evaluation. However, they decided not to change their direct evaluation and used the inconsistencies as discussion points to be shared with the project managers as dialogue in future team meetings. Recognizing these inconsistencies and talking about them improved the participant’s and the team managers’ discourse.

Participants who made changes to their direct and indirect evaluations were either biased and thinking about criteria that were not contained in the evaluation table or improving their learning about tradeoffs. For instance, Participant 13 changed their direct evaluation after recognizing bias for some of the strategies (e.g., preferences for strategies that may be more easily implemented in their home state) and criteria that were not included in the evaluation table (e.g., funding during the COVID-19 pandemic). This improved their consistency by three orders of magnitude (Figure 3, upper panels).

Participants 1, 6, and 12 adjusted their swing weights to improve consistency. In all instances, the intensity of
preference for tradeoffs in the evaluation table were not fully expressed in their initial swing weighting. Each of these participants provided an even distribution of weights in their initial indirect evaluation (Figure S1). In general, they assigned high point values to the swings. This occurrence is typically associated with ascribing too much importance to less-important criteria and has been documented in some experiments (Shepard, 1964; Von Winterfeldt & Edwards, 1986). The occurrence can signify a lack of cognitive strength in using numbers to express a highly subjective state of mind; this is a limitation of any approximate or surrogate weighting method.

**FIGURE 1** Final cumulative weights per participant and weight distribution across participants. Points and bars reflect individual participant weights and range over all participant weights, respectively. Box and whisker plot contains 50% interquartile range, separated by the median line, and whiskers contain 90% interquartile range with outlier points.
Final rankings from direct (x-axis) and indirect (y-axis) methods. Dashed centerline represents consistency. Strategies that fall on the centerline rank consistently between methods, whereas strategies that fall further from the centerline rank inconsistently between methods.
Through deliberation, however, we found that some weight distributions should have been more unevenly spread based on reasons associated with their direct evaluation. For instance, Participant 12 spread the emphasis on feasibility and risk as less-important objectives than impact. One of the reasons for their adjustment was that they had a difficult time assigning preferences to the objectives, represented in the swing weighting as bundles of inter-objective criteria performance values in a hierarchical sense, even though they had strong preferences for some criteria more than others. This adjustment improved their consistency by four orders of magnitude (Figure 3, lower panels). In this and other instances, spreading the point intervals across the swings spread out the weight distribution and improved consistency between the methods. Each of these participants adjusted their indirect rankings based on reflection and deliberation and subsequently felt better about their final rankings.

This study encountered relief and stress with navigating inconsistencies, but it is unclear whether these feelings are determinants to inconsistencies between methods because no structured survey was included in the evaluation. For instance, Participant 7 understood the process better after adjusting their direct and indirect evaluations. This participant (and others in this study) recognized response error in their indirect evaluation, which is a common phenomenon (Jia, Fischer, & Dyer, 1998), and subsequently improved their consistency by four orders of magnitude. Changing preferences
over time is not a problem, and this participant exhibited relief in a tool that allowed them to think about and adjust their preferences. On the other hand, Participants 9 and 5 found the indirect method more stressful than the direct method. Participant 9 changed both direct and indirect evaluations twice, and they were the only participant that did not improve consistency relative to their initial rankings. Participant 5 was the most inconsistent between methods (Figure 2), which was due in part to highly variable criteria weights (Figure 1) and potential bias in their direct evaluation. Considering it hard to quantify preferences, these participants gained the ability to distinguish tradeoffs after discussing their inconsistencies with the author.

The tradeoff in stress levels between the direct and indirect methods is consistent with previous experimental and applied studies (Arvai & Gregory, 2003; Arvai, Gregory, & McDaniels, 2001; Gregory et al., 2012; Hobbs & Horn, 1997). A goal of the indirect method is to make tradeoffs explicit, which may be hard for participants to intuit. Contrary to the ease of implementing the direct method, participants generally found it hard to disentangle technical information on the strategies into a consistent preference structure. Likewise, focusing solely on the strategies revealed bias in various forms. Combining the swing weighting with follow-up discussions helped to safeguard some but not all confusion associated with the indirect method. In summary, the tradeoff is not always straightforward and can be alleviated to increase the chances that a multimethod approach is insightful for decision making.

This study found that bias, incomplete criteria, and challenges with criteria weighting were the sources of inconsistencies between direct and indirect evaluations (Table 3). These findings are consistent with previous studies (Bond et al., 2008; Gregory et al., 2012; Keeney et al., 1990). The multimethod approach helped to assist participants in confronting their inconsistencies and reveal areas of agreement and disagreement in their preferences. Explicit discussion about tradeoffs using the indirect method allowed participants to consider a wide variety of decision-relevant issues that worked in parallel with the direct method (Arvai & Gregory, 2003). Most of the participants noted that these discussions increased their confidence in the results. This learning aspect of the multimethod approach can improve conservation science and practice.

5 | CONCLUSIONS

This study used decision-analytic concepts and tools to understand inconsistencies in team member preferences and prioritize The Nature Conservancy’s conservation strategies in agricultural landscapes of the Chesapeake Bay watershed. Final decisions on the strategies were consistent with the mean final ranking across participants (Figure 2, last panel). Project managers decided to advance three strategies that had the highest team ranking and consistency between direct and indirect evaluations. One strategy that had the lowest team ranking and consistency will not be pursued at this time. Two strategies that had inconsistent ranking among team members require further clarification and research to better understand some of their assumptions. The project managers noted the strength and enhanced comfort level of comparing participants’ direct evaluations with indirect evaluations to make the final decisions. One of the project managers shared an added degree of confidence in results from the indirect method compared with the direct based on its ability to reveal preferences for tradeoffs and reduce bias.

It is important to note that a multimethod approach does not guarantee consensus, especially in collaborative group decision processes. Decision analysis should not be used in place of a final decision, but only to improve insights into the desirability of decision alternatives. In addition, the approach takes more time to implement than any single method. This limitation is counterbalanced by more quality information that decision makers can use to make informed choices. There are no hard-and-fast rules about when a multimethod approach is appropriate to analyze a decision. There is some indication that the reliability of such techniques decreases with increasing problem dimensions (e.g., number of alternatives or objectives/criteria; Von Winterfeldt & Edwards, 1986). Likewise, decision context and decision-maker intentions can be used to indicate the necessity for the approach.
The methods were greatly improved with the aid of a decision facilitator. Facilitators work directly with decision makers to ensure that any analysis is implemented with the highest scientific standards and provides useful insights, and they also work directly with participants to ensure that all aspects of the decision are considered (Phillips & Phillips, 1993). Among the many roles of a decision facilitator are making everyone aware of analytical assumptions, assuring that everyone understands the methods and results, and controlling for bias, where relevant (Salo & Hämäläinen, 2010).

This study presents findings of an internal organizational decision-making process. Further research on scaling and integrating these methods among conservation practitioners and multiple partners is important. The various approaches to conservation planning (e.g., Conservation Measures Partnership, 2020; The Nature Conservancy, 2016) provide a framework to partner with outside organizations and uniformly achieve conservation goals. This framework is embraced by conservation organizations around the world. This study provides context to improve upon this framework with further case studies and practice.

ACKNOWLEDGMENTS
Special thanks to Michael C. Runge, Ralph L. Keeney, Amy Jacobs, and Mark Bryer for feedback on methods, The Nature Conservancy’s Chesapeake Bay agriculture team for participating in the research, and Joshua H. Goldstein, Amy Jacobs, and anonymous journal referees for manuscript reviews.

CONFLICT OF INTEREST
The authors declare no conflicts of interest.

AUTHOR CONTRIBUTION
David M. Martin: Designed and conducted the research, carried out the data collection and interpretation, and wrote and edited the manuscript.

DATA AVAILABILITY STATEMENT
Supporting data are available on request: david.martin@tnc.org.

ETHICS STATEMENT
The author is not aware of any ethical issues regarding this work. Research involving human subjects was approved by The Nature Conservancy, Office of the Chief Scientist.

ORCID
David M. Martin https://orcid.org/0000-0002-1514-5734

ENDNOTES
1 For more information: https://www.nature.org/en-us/about-us/where-we-work/united-states/maryland-dc/stories-in-maryland-dc/mddc-how-we-work-agriculture/.

REFERENCES
Arvai, J., & Gregory, R. (2003). Testing alternative decision approaches for identifying cleanup priorities at contaminated sites. Environmental Science & Technology, 37, 1469–1476. https://doi.org/10.1021/es020762p
Arvai, J. L., Gregory, R., & McDaniels, T. L. (2001). Testing a structured decision approach: Value-focused thinking for deliberative risk communication. Risk Analysis, 21, 1065–1076. https://doi.org/10.1111/0272-4332.216175
Bessette, D. L., Wilson, R. S., & Arvai, J. L. (2019). Do people disagree with themselves? Exploring the internal consistency of complex, unfamiliar, and risky decisions. Journal of Risk Research, 1–13. https://doi.org/10.1080/13669877.2019.1569107
Bessette, D., Wilson, R., Beauvirie, C., & Schroeder, C. (2019). An online decision support tool to evaluate ecological weed management strategies. Weed Science, 67, 463–473. https://doi.org/10.1017/wsc.2019.21
Bond, S. D., Carlson, K. A., & Keeney, R. L. (2008). Generating objectives: Can decision maker articulate what they want? Management Science, 54, 56–70. https://doi.org/10.1287/mnsc.1070.0754
Chesapeake Assessment and Scenario Tool. (2020). Source data. Retrieved from https://cast.chesapeakebay.net/
Chesapeake Bay Commission. (2019). Annual report. Retrieved from https://www.chesbay.us/annual-reports
Conservation Measures Partnership. (2020). Open standards for the practice of conservation, v.4.0. Bethesda, MD: Conservation Measures Partnership. Retrieved from https://conservationstandards.org/
Danielson, M., & Ekenberg, L. (2016). A robustness study of state-of-the-art surrogate weights for MCDM. Group Decision Negotiation, 26, 677–691. https://doi.org/10.1007/s10726-016-9494-6
Dyer, J. S., & Sarin, R. (1979). Measurable multiattribute value functions. Operations Research, 27, 810–822. https://doi.org/10.1287/opre.27.4.810
Fischhoff, B., & Davis, A. L. (2014). Communicating scientific uncertainty. Proceedings of the National Academy of Sciences, 111(4), 13664–13671. https://doi.org/10.1073/pnas.1317504111
Foundations of Success. (2018). Conceptualizing and planning conservation projects and programs: A training manual. Bethesda, MD. Retrieved from: Foundations of Success. https://fosonline.org/library/conceptualizing-and-planning-manual/
Gardiner, P. C., & Edwards, W. (1975). Public values: Multiattribute utility measurement for social decision making. In M. F. Kaplan & S. Schwartz (Eds.), Human Judgement and decision processes. New York, NY: Academic Press.
Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., & Ohlson, D. (2012). Structured decision making: A practical guide to environmental management choices. West Sussex, United Kingdom: Wiley-Blackwell.

Hobbs, B. F., & Horn, G. T. F. (1997). Building public confidence in energy planning: A multimethod MCDM approach to demand-side planning at BC gas. Energy Policy, 25, 357–375. https://doi.org/10.1016/S0301-4215(97)00025-6

Hostmann, M., Bernauer, T., Mosler, H.-J., Reichert, P., & Truffer, B. (2005). Multiattribute value theory as a framework for conflict resolution in river rehabilitation. Journal of Multi-Criteria Decision Analysis, 13, 91–102. https://doi.org/10.1002/mcda.375

Jia, J., Fischer, G. W., & Dyer, J. S. (1998). Attribute weighting methods and decision quality in the presence of response error: A simulation study. Journal of Behavioral Decision Making, 11, 85–105.

Kahneman, D. (2013). Thinking, fast and slow. New York, NY: Farrar, Straus, and Giroux.

Keeney, R. L. (1992). Value-focused thinking. Cambridge: Harvard University Press.

Keeney, R. L. (2004). Framing public policy decisions. International Journal of Technology, Policy and Management, 4, 95–115. https://doi.org/10.1504/IJTPM.2004.004815

Keeney, R. L. (2020). Give yourself a nudge: New ways to make better personal and business decisions. New York, NY: Cambridge University Press.

Keeney, R. L., & Gregory, R. S. (2005). Selecting attributes to measure the achievement of objectives. Operations Research, 53, 1–11. https://doi.org/10.1287/opre.1040.0158

Keeney, R. L., Von Winterfeldt, D., & Eppel, T. (1990). Eliciting public values for complex policy decisions. Management Science, 36, 1011–1030. https://doi.org/10.1287/mnsc.36.9.1011

Martin, D. M., & Mazzotta, M. (2018). Non-monetary valuation using multi-criteria decision analysis: Sensitivity of additive aggregation methods to scaling and compensation assumptions. Ecosystem Services, 29, 13–22. https://doi.org/10.1016/j.ecoser.2017.10.022

Martin, D. M., Mazzotta, M., & Bousquin, J. (2018). Combining ecosystem services assessment with structured decision making to support ecological restoration planning. Environmental Management, 62, 608–618. https://doi.org/10.1007/s00267-018-1038-1

McDaniels, T. (2019). Four decades of transformation in decision analytic practice for societal risk management. Risk Analysis, 39, 859–868. https://doi.org/10.1111/risa.13332

Phillips, L. D., & Phillips, M. C. (1993). Facilitated work groups: Theory and practice. The Journal of the Operational Research Society, 44, 533–549. https://doi.org/10.2307/2584511

Salo, A., & Hämäläinen, R. P. (2010). Multicriteria decision analysis in group decision processes. In D. M. Kilgour & C. Eden (Eds.), Handbook of group decision and negotiation. Dordrecht, Netherlands: Springer.

Shepard, R. N. (1964). On subjectively optimum selection among multiattribute alternatives. In M. W. Shelley & G. L. Bryan (Eds.), Human Judgement and optimality. New York, NY: Wiley.

Slovic, P. (1995). The construction of preference. American Psychologist, 50, 364–371. https://doi.org/10.1037/0003-066X.50.5.364

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1976). Cognitive processes and societal risk taking. In H. Jungermann & G. de Zeeuw (Eds.), Decision making and change in human affairs. Dordrecht, Netherlands: Springer.

Straus, D. (2002). How to make collaboration work. San Francisco, CA: Berrett-Koehler Publishers Inc.

The Nature Conservancy. (2000). The five-S framework for site conservation: A Practitioner’s handbook for site conservation planning and measuring conservation success. Arlington, VA: The Nature Conservancy. Retrieved from http://d2ouvy59p0dg6k.cloudfront.net/downloads/5S-v1-1.pdf

The Nature Conservancy. (2016). Conservation by design 2.0. Retrieved from https://www.conservationgateway.org/ConservationPlanning/cbd/Pages/default.aspx

U.S. Environmental Protection Agency. (2010). Chesapeake bay total maximum daily load for nitrogen, phosphorus, and sediment. Philadelphia, PA: US Environmental Protection Agency, Water Protection Division.

Von Winterfeldt, D. (2013). Bridging the gap between science and decision making. Proceedings of the National Academy of Sciences U S A, 110, 14055–14061. https://doi.org/10.1073/pnas.1213532110

Von Winterfeldt, D., & Edwards, W. (1986). Decision analysis and behavioral research. New York, NY: Cambridge University Press.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Martin, D. M. (2021). Navigating inconsistent preferences: A multimethod approach to making informed decisions. Conservation Science and Practice, e469. https://doi.org/10.1111/csp2.469