International Centre for the Study of
the Preservation and Restoration of
Cultural Property

ICCROM

Sharing Conservation Decisions

Current Issues and Future Strategies

Edited by

Alison Heritage & Jennifer Copithorne

ICCROM

Published by the International Centre for the Study of the Preservation and Restoration of Cultural Property (ICCROM), Via di San Michele 13, 00153 Rome, Italy.

ISBN 978-92-9077-271-2

Sharing Conservation Decisions: Current Issues and Future Strategies
Edited by Alison Heritage & Jennifer Copithorne

Cover artwork: Karen Abend
Typesetting: Datapage International Limited

Coordination: Catherine Antomarchi, Jennifer Copithorne, Isabelle De Brisis, Alison Heritage & Isabelle Verger

Sharing Conservation Decisions: Tools, Tactics, and Ideas

STEFAN MICHALSKI

ABSTRACT

Studies of decision sharing in heritage conservation show that sharing, even when attempted, usually fails to influence the decision. A far larger study of decisions made in commercial and non-commercial organizations showed that the failure to share the decision was a major cause of poor outcomes. Two common decision-making tools, the decision matrix and the decision tree, are explained. Conservation examples are discussed, including the decision matrix developed during the case study of the ICCROM Sharing Conservation Decisions (SCD) 2008 course. Tactics and ideas for effective sharing of decisions are drawn from recent texts on participatory decision-making, cognitive psychology, and experimental moral philosophy. The ability to be reflective, identified by Stanovich as a separate trait from intelligence, emerges again and again as the key piece of advice from the decision-making literature. The author concludes that, for our field, decision tools are best understood as a means to structure and document shared reflection, not to automate what are always difficult decisions.

Introduction

The literature on decision-making spreads across several disciplines, from business management to mathematics to psychology to philosophy. As documented by Antomarchi and Abend (2017) elsewhere in this volume, interest in the topic has also grown rapidly in our own field, presumably out of necessity rather than idle curiosity.

For almost four decades I have been a technical expert to museums and galleries during their decision-making about environmental control and lighting. For the Dahlem Conference of 1992, I explored far outside my technical expertise to see what the fields of perception, structuralism, and museology might teach us about "Sharing responsibility for conservation decisions" (Michalski, 1994). Over the next two decades, I developed tools for a particular form of quantitative decision-making – risk assessment – which led me to scan much of the risk and decision literature. For the SCD course in 2008, I volunteered to provide an introduction to these readings, and to guide an exercise in the application of a standard decision-making tool – the decision matrix. In 2010, I collaborated with a painting conservator to apply a second standard tool, the decision tree, to document not only the reasoning behind the final treatment, but also to document the many treatment options that had been considered but rejected (Michalski and Rossi-Doria, 2011). By the time I sat down to revise my 2008 notes for this article in 2016, excellent texts covering the same ground had been written by experts on decision-making (Kahneman, 2011; Manktelow, 2012), experts on moral decisions (Greene, 2013; Haidt, 2013) and experts on facilitation of participatory decision-making (Kaner, 2014; Renn, 2015). And just six months earlier, our own field produced an overview of the literature of decision-making, with recommendations for conservators (Henderson and Waller, 2016).

Rather than attempt yet another overview (which tend to leave the reader pessimistic about whether they can make good decisions without years of preparation), I have focused on practical advice from three sources: 1) published evidence about the rate of success in sharing decisions; 2) basic decision-making tools that are widely promoted for managers in general; and 3) recent researches into the way we humans think about these issues.

Published by ICCROM.

Evidence of failure in the sharing of decisions

Failure in sharing decisions on the treatment of movable heritage

For years, gadflies on our profession's web forums have been begging our field to report honestly on the aftermath of conservation decisions. We are not alone in this failing – the health field, thousands of times bigger than us, only began to do this kind of 'post-mortem' on trendy treatments in the late 1980s and early 1990s. It became known as 'evidence-based policy' and then as 'evidence based medicine' (see https://en.wikipedia.org/wiki/Evidence-based_medicine).

Henderson and Nakamoto (2016) examined 32 published case studies of conservation projects that consulted with stakeholders of some sort. They separated the sharing processes within the case studies into three types: sharing during appraisal of the meaning and context of the objects; sharing when deciding about treatment; and sharing when deciding about display or storage. They assigned each case study to one of two categories – those where the stakeholder advice was 'ignored', and those where it was 'acted upon.'

Most of the projects shared the appraisal stage, and advice received was never ignored – not surprising since there is no decision to dispute, only a neutral pooling of knowledge. Over half the projects shared decisions about the third stage – display and storage – and in none of these was stakeholder advice ignored. While attempting to share treatment decisions, however – the stage near and dear to the conservator's heart – stakeholder influence collapsed. To begin with, less than half the projects that solicited stakeholders even considered sharing the treatment decision, and of those that did, half ignored the advice anyway. If we only consider non-expert stakeholders, ignoring rose to three-quarters of the attempts (Table 1).

This analysis by Henderson and Nakamoto (2016) does not allow one to draw direct conclusions about the rate of failure to share decisions *per se*. It does, however, establish that, even when prepared to share parts of their project (and write about it), conservators will usually not consider sharing decisions about their special area of competence, and if they do, they will then ignore that advice most of the time.

Failure in sharing decisions in the management of immovable heritage

In an editorial in an issue of *The International Journal of Heritage Studies* devoted to community engagement in site management, Watson and Waterton refer to "box ticking expediencies associated with ideas about social-inclusiveness" and "a kind of self-satisfaction in the heritage community that the job had, indeed, been done" (2010, p. 1). Inside the issue, Chirikure *et al.* examined three world heritage sites in Africa and found "many professionals pay lip-service to the whole concept of participation because the interests of the local communities and those of professionals do not always coincide" (2010, p. 30). While not a systematic study of a large number of cases that would constitute evidence based conservation in its strict sense, it is an honest examination of what actually happened despite good intentions and important projects. The authors "feel strongly that there is a need for active research programmes by heritage managers to generate information for management as well as for empowering local communities" (2010, p. 41). This is a far more radical proposal than simply asking experts to listen; it asks experts to create and hand over knowledge, i.e. power, to the community, who will then make decisions and sustain the project. This goes further than Renn's observation, after examining successful risk assessment consultations in the field of public safety, that "participants from the lay public were not only willing to accept, but furthermore demanded that the best technical estimate of the risks under discussion should be employed for the decision-making process" (2008, p. 330).

Evidence of failure in sharing decisions by managers in general

Nutt (2002) examined over 400 major management decisions spanning 20 years, made in businesses, non-profits, and public organizations. He examined the methods used and the eventual outcomes. He concluded that "half the decisions made in organizations fail". He also examined well-known "debacles" to see what went wrong. Nutt discovered three fundamental "blunders", each of which contain failures to share.

The first blunder Nutt called "the rush to judgement". Managers identified a concern and latched on to the first remedy that they came across, especially when those higher up pressured them. The rush to judgement caused failures four times more often than when managers took the time to investigate thoroughly. Investigating thoroughly generally means sharing the decision and its context with others.

The second blunder was "misuse of resources". Managers spent their time and money during decision-making on the wrong things, for example spending heavily on evaluations in attempts to defend the first type of blunder, rather than gathering useful information in the first place.

The third blunder was the use of "failure-prone tactics" – used in two-thirds of all decisions. For example, although managers knew that sharing with staff was important, they used it only 20 percent of the time! The data showed that staff participation resulted in an 80 percent success rate. Another failure-prone tactic was the use of coercion by managers, applied in 60 percent of the decisions, but successful only 30 percent of the time. Coercion of staff is the opposite of sharing with staff!

Managers who made one of Nutt's three blunders found themselves caught in one or more of seven traps: (1) failing to uncover concerns and competing claims (not sharing); (2) overlooking people's interests and commitments (not sharing); (3) leaving expectations vague (not sharing); (4) limiting the search for options; (5) misusing evaluations (not sharing); (6) ignoring ethical questions (not sharing); and finally, (7) failing to reflect on earlier results to learn what worked and what did not.

If one examines all these blunders and many of the traps, drawn from a massive collection of evidence, one sees that one way or another, they are all failures to share the decision, whether with experts (which takes time, money, and effort) or with stakeholders (which requires one to abandon coercion, consider people's interests, and address ethics).

Nutt also distinguished between decisions made with an "idea-driven" process, which defined the problem or its possible options (the idea) very early in the process, versus a "discovery-driven" process that took the time to explore the actual definition of the problem as well as the options. He found poor outcomes were four times more likely with the idea-driven process, and that all the "debacles" had used the idea-driven process. A key difference between the two processes is the early and honest sharing not only of the decision, but its formulation in the first place, i.e. the goal. Perhaps, for example, stakeholders invited to select among predetermined treatment or exhibition options are not interested in those decisions at all, rather they want to decide which objects or parts of a site to consider in the first place.

Tool 1: the decision matrix

A brief history of the decision matrix

Benjamin Franklin proposed a method for decision-making, based on a list of pros and cons, which one then crossed off in pairs, taking account of their relative weight, until only one side remained (Yoon and Hwang, 1995). By the eighteenth century, the utilitarian philosopher, Bentham argued that only a moral arithmetic, the summing of the greater good, could decide whether actions were moral or not (Driver, 2014). Such utilitarian logic still underlies the preservation and access goals of conservation decisions today (Michalski, 2008). Greene (2013) argues convincingly that it remains the only rational principle for moral decision-making in general.

The multicriteria decision matrix emerged in its current form (Table 2, Figure 1) for prosaic business decisions in the 1950s, and by 1968 it was an established method applied in almost a hundred different journals (Hwang and Yoon, 1981). A decade later, Kepner and Tregoe (1976) promoted the tool in their book *The rational manager; a systematic approach to problem solving and decision-making*. Today, the literature extends to highly mathematical theories where hundreds of options, criteria, and probabilities are in play (Hwang and Yoon, 1981; Yoon and Hwang, 1995), but the average manager, if they use a decision matrix at all, uses the same simple types of fifty years ago (Mindtools,2017).

	Number of projects examined	Number of projects that shared the Appraisal Stage and percentage where advice ignored		Number of projects that shared the Treatment Stage and percentage where advice ignored		Number of projects that shared the Display and Storage Stage and percentage where advice ignored	
TOTALS	32	26	0%	13	46%	18	0%
Museum professionals	10	8	0%	5	100%	4	0%
Religious community	6	6	0%	3	67%	2	0%
Community of origin	10	7	0%	4	75%	8	0%
Artists	6	5	0%	1	100%	4	0%

Table 1. The proportion of case studies where stakeholder advice was ignored, as found by Henderson and Nakamoto (2016) in an analysis of 32 published conservation projects.

The arithmetic of the decision matrix: adding up good points

Table 2 illustrates a decision matrix applied to a very common conservation decision – choosing between imperfect treatment options for flaking and powdery paint, each imperfect in its own way. The scores and weights of the original case study (Michalski and Rossi-Doria, 2011) have been adjusted so as to illustrate better issues that are discussed below.

The rows of Table 2 contain four criteria – reversibility, stability, appearance, and speed of application. (These are almost universal in a conservator's judgement of treatments.) The specific definitions used were as follows: 'appearance' means the appearance immediately after treatment; 'stability' means primarily the estimated change in appearance after 100 years and 'speed' refers to the total labour cost. Under 'stability', the threshold of minimally acceptable degree of yellowing is defined as; noticeable but not disfiguring after 100 years (best available estimates), and this is assigned a score of three.

As is usually the case, stability versus appearance presents a trade-off: Treatment A has excellent stability (stable polymers, 5 out of 5), good speed (4 out of 5), but poor appearance (1 out of 5, it darkens the object noticeably). Treatment C, a traditional method, is the complete reverse – looks great today (5 out of 5) and applies easily (5 out of 5), but is predicted to be very yellow in much less than 100 years (1 out of 5). Treatment B scores well on appearance and stability but is extremely laborious (application of consolidant flake by flake). If the decision-makers had decided that there was to be no mandatory minimum on stability, then Treatment C would emerge as the best option (11 points), but given the minimum acceptable stability of three points, then Treatment A emerges as the best option before weighting (10 points).

Weighting: some issues are more important

It is unusual for criteria to be equally important. One can correct this imbalance by assigning different 'weights' to each criterion. In Table 2, the appearance has been weighted as most important: weight 3.

Stability (which determines future appearance) has been weighted almost as important: weight 2 (plus it has a minimum). Speed is considered least important. Perhaps these are the weightings of a major museum rather than a private client! The decision now shifts to Treatment B.

The overall lesson from these switches in decision from Treatment C to A to B is not that one can 'play' the matrix to get what one wants, but rather that the matrix can capture the reasons that a decision might shift between plausible options. In this case, setting a minimum stability means that an otherwise excellent treatment is rejected, and deciding that speed (cost) is much less important will shift the decision yet again. In other words, the tool documents the individual judgements that have been considered, documents the judgement about the relative significance of those judgements, and then points to the decision consistent with those judgements.

The decision matrix of the case study of SCD 2008

Figure 1 presents a screen capture of the decision matrix compiled by participants of the SCD 2008 course for their case study. A spreadsheet as shown can easily be created for a decision matrix by anyone familiar with basic formulae in Excel™. The case study was a religious site with multiple buildings. The municipality wanted a long-term plan that satisfied many different users – religious pilgrims, tourists, locals – as well as its economic realities as the custodian. (Unlike the previous example in Table 2, the scores in Figure 1 have not been modified for didactic purposes.)

After much discussion, the course participants decided on the criteria shown, and votes were taken to establish the weightings (group averages). A scale of 1 to 9, rather than 1 to 5, was used for both weightings and scoring, as recommended by various authors to allow smaller differences to emerge. The weightings voted by the group are entered in column W1.

The four options in Figure 1 were contained in detailed reports developed by four working groups over many days. The scores were voted on by the course participants who were not in the design

	Options:	1		2		3		4		W: weights			select weight group
	Criteria	score	x W	score	x W	score	x W	score	x W	W 1	W 2	W 3	
#1	Low impact on material and visual integrity in 5 years	7.4	46	6.8	42	7.6	46.9	7.3	45.2	6.2	6.2	6.2	1
#2	Low impact on material and visual integrity in 30 years	6.8	44	6.2	40	6.0	39.0	5.9	38.1	6.5	6.5	6.5	
#3	Low impact on current religious use of site	4.9	32	4.8	32	7.7	50.9	6.1	40.5	6.6	9.0	5.8	
#4	Benefits to community in 5 years	7.1	41	6.8	39	6.4	37.3	6.6	38.1	5.8	5.8	5.8	
#5	Expected increase in cultural tourism	6.6	43	6.9	45	6.6	42.7	7.7	50.1	6.5	6.5	6.5	
#6	Sustainability of the managing institution	6.8	41	6.0	37	5.4	33.1	5.4	33.1	6.1	6.1	6.1	
	Totals without weights	39.6		37.4		39.7		39.0					
	Totals with weights		248		235		250		245				
	Percent of maximum score		99%		94%		100%		98%				

Figure 1. A decision matrix made in Excel™ as used for the case study of SCD 2008. It highlights the best scores in each criterion in green, and allows three different sets of weightings to be entered and compared easily.

		Treatment A		Treatment B		Treatment C	
CRITERIA	WEIGHT	score	x weight	score	x weight	score	x weight
Reversible (must be)	-	PASS		PASS		PASS	
Appearance	3	1	3	4	12	5	15
Stability (must be 3+)	2	5	10	4	8	1 FAIL	2
Speed	1	4	4	1	1	5	5
Total score		10	17	9	21	11	22
Comments		Best stability, good speed, but poor appearance		Good appearance, good stability, but poor speed.		Best speed, best appearance, but stability FAIL.	

Table 2. Example of a simple decision matrix with scores on a five-point scale.

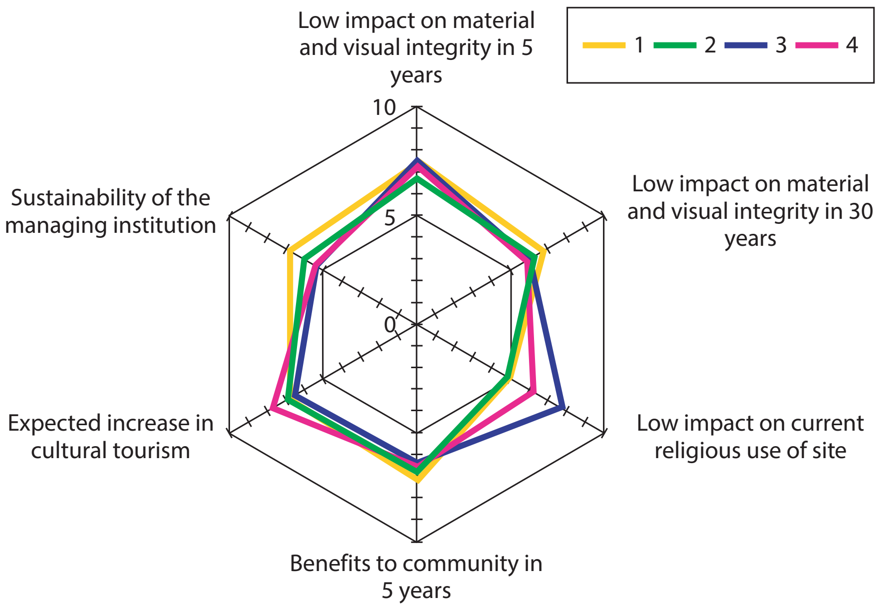

Figure 2. A radar chart of the four options shown in Figure 1.

group, after presentations of the proposals. The overwhelming conclusion from the numbers in Figure 1 is that there was little difference in the weightings of each criterion, and little difference in the totals of the four options. This does not mean that the decision matrix was useless; it simply meant that all four options were well designed, although distinctly different. The conclusion of the class after this first iteration of option and criteria development was that in the real world, one would want to take the lessons learned and build even better options and better criteria with better representation of stakeholders before making a final decision.

Radar chart

Figure 2 presents a 'radar chart' of the options in Figure 1. (Also known as a spider chart, web chart, or star chart.) Radar charts are standard in Excel™ and many other graphing tools. One plots the unweighted scores to see how well options perform across various criteria. In Figure 2, we can see that on most criteria the options score similarly, that none fall below a score of five, but that the biggest differences occur on the lower right axis – impact on current religious use of the site.

Sensitivity analysis

In Figure 1, columns W2 and W3 are used as ‘what if’ weightings that can be quickly changed to see how the decision changes. Selecting the weights of column W2 will show what happens if the criterion ‘Low impact on religious use of the site’ is given a maximum weighting (nine points). It was found that the decision did not change, it remained on option 3. For the weights of column W3, the weighting for religious use (criteria #3) was lowered until the decision changed. It was found that down to a weight of 5.8, the decision was unchanged, but at a weight of 5.7 or less the decision shifted to option 1. It can be very helpful to building consensus if one can demonstrate that a decision is not sensitive to the range of opinions on a particular weighting or score. In this example, weighting of the religious use criteria can range from 5.8 to 9 without the decision changing.

Software tools

A matrix that does the arithmetic behind the scenes can easily be built with rudimentary knowledge of formulas in any spreadsheet software. One can find free decision matrix templates online that use Excel™. Features of the spreadsheet created by the author for the SCD 2008 course, Figure 1, include quick toggling between three different sets of weights, and conditional formatting to highlight the option with the best scores on each criterion (green cells, Figure 1).

This article does not survey specific decision-making gadgets and software one can find online – they come and go too quickly. Free tools tend to keep your data online, tools that stand alone on your computer tend to be expensive. That said, there are some online tools that facilitate the process of weighting criteria by using ‘pairwise comparisons’ and sliders that make selection of scores more visual.

When the goal emerges after the criteria

As Henderson and Waller (2016) stress, one should clarify one’s goal before setting up any decision-making process. In risk management, for example, it might be “to minimize expected loss of asset value as measured 100 years in the future”. For many decisions however, definition of a goal before defining criteria is not so simple. The classic example given in texts on decision-making is that of someone selecting a car (or now a smart-phone). The criteria are often contradictory – initial cost, fuel efficiency, prestige, sportiness, cargo capacity. The most common expression of the goal for such decisions is simply ‘the best all-around option’ whether car or conservation treatment. The key to understanding whether the selected criteria will constitute the correct goal is to understand *for whom* we are seeking this 'best' option and whether they agree with the criteria and their weightings. Sharing decisions is not simply about incorporating the knowledge of others, but also about accepting the utilitarian ethic that we are trying to maximize the greater good, and that we can only determine that by understanding the consequences for all those affected.

One technical aspect of goal setting that does have universal applicability to conservation decisions (and most business decisions) is the time horizon. Do you want the best decision as judged in terms of one year, 10 years, 100 years, or longer. This has been incorporated into the stability criteria of Table 2, and almost every criteria of the example in Figure 1. This is an expansion of the utilitarian perspective to sharing the decision with future generations.

Musts

Kepner and Tregoe (1976) advised that one can set some criteria to a 'must'. When the criterion can be answered with a pass/fail, quantification is no longer an issue. In conservation treatment decisions such as Table 2, reversibility is usually set as a 'must' (even though we all know it is never so simple). We sometimes neglect to consider a 'must' because it is presumed, but a decision matrix should consider 'musts' explicitly, enabling their re-examination if they block a shared decision.

A second 'must' in Table 2 is stability, this time expressed as a minimum acceptable degree of change in 100 years. Minima need a measure of some kind in order to be usable and negotiable.

Making a decision based only on 'musts' is known as "conjunctive satisficing" (Hwang and Yoon, 1995; Manktelow, 2012). One accepts any option that meets a set of 'musts'. In Table 2, Treatments A and B satisfy all 'musts' (reversibility and minimum stability). At that point, one can just flip a coin, or engage with the arithmetic of the decision matrix to identify the best between Treatments A and B.

Building an ensemble of different strengths

If one sets very high minima for all criteria and accepts that no single option will meet all of them, one can decide to accept options that meet some of them. This is called disjunctive satisficing (Yoon and Hwang, 1995; Manktelow, 2012). This approach emerges when each decision is part of a larger process. For example, when building a team of experts, one might accept an expert that meets some of the stringent criteria. The next expert must then satisfy some of the remaining criteria, and so on. In Hedley's (1990) discussion of the three options (schools) for the cleaning of paintings, he proposes that the only criterion for which all options should meet a high minimum is competent implementation of their particular school of cleaning. All other criteria, such as respect for original materials, recovery of artistic intention, aesthetic integrity, respect for object history, minimal intervention, etc. will be met very differently by the different schools. If we made a radar plot using all the competing criteria and placed all the schools proposals on it, we would find plots that all shared a high score on competence but otherwise scored well only on those criteria favoured by the goal of each school. Hedley concludes that this global ensemble provides a richer result for humanity than a single standard.

Consider the standard approach to the display of light-sensitive objects: rotation of the collection. We can agree that in a perfect world, we would set very high minima for both access and preservation, i.e., criteria 1 is "objects fade negligibly over centuries" and criteria 2 is "objects are seen well every day by visitors". Rotation fails both, it presumes that one must lower the minima of both criteria and find a conjunctive solution, i.e. a compromise. Disjunctive reasoning would look for paths that have part of the collection meeting the difficult minimum of criteria 1, so that part must stay archived in storage, and the other part of the collection must score well on criteria 2, so that part is on display permanently. This is an ensemble or teamwork solution – there will always be authentic brightly coloured exemplars available for whatever new reproduction technology comes along.

Tool 2: the decision tree

Although the name 'tree' for diagrams such as Figure 3 was inevitable, I believe it is the metaphor of paths taken and paths not taken that helps to explain the power of decision trees. There are two varieties of decision tree: predicting a set of outcomes, and guiding a sequence of contingent decisions.

Decision trees that calculate a set of outcomes usually incorporate probabilities of success along each path from each node. These trees begin on the left side of the page with an initial entry point, and end on the right side with a long column of possible end results that are the product of the interacting probabilities. Caple (2000) provides two examples for a conservation manager exploring collection care options in terms of costs and benefits.

Figure 3 is a decision tree for a range of possibilities in treating a painting. On the right-hand side, the predicted outcome of each possibility is given a score using pluses and minuses. The purpose of this tree was not to make the decision, but to document the many possibilities that were carefully considered but rejected (Michalski and Rossi-Doria, 2011). This tree also incorporates a small decision matrix at the end of the dominant pathway. (Trees and matrices are not incompatible.)

Decision trees that guide a sequence of smaller decisions look exactly like Figure 3, with simple yes/no decisions directing one's path, but rather than using the many endpoints to determine the best path of all, these decision trees point you down the right path for your

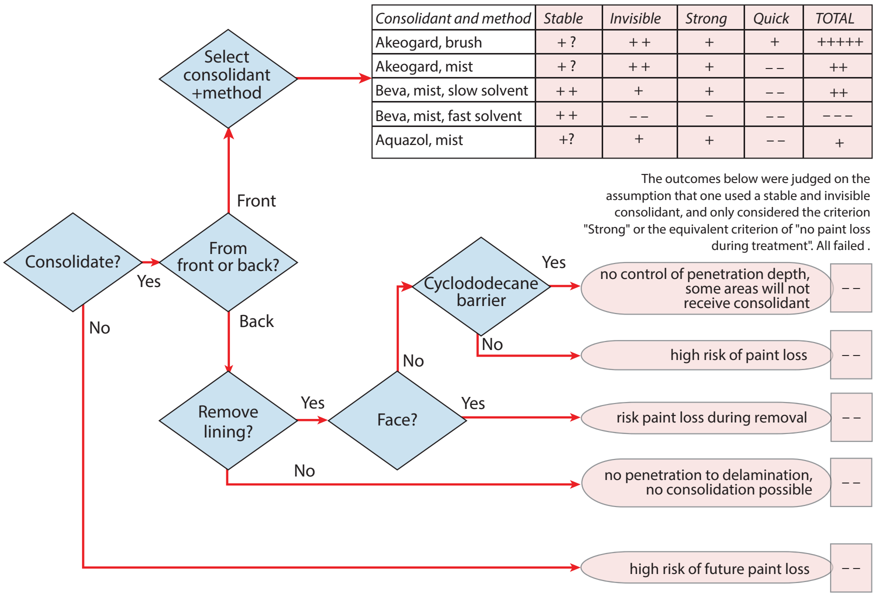

Figure 3. A decision tree for comparing various treatment options for a painting (Michalski and Rossi-Doria, 2011).

particular situation. Strang (2003) provides an example in our field for the processing of electronic records on arrival at an archive (itself a summary of a much larger tree used by large archives).

Tactics for sharing

Sharing what with whom

The authors cited earlier on the failure of sharing decisions all note that defining the groups with which one will share is essential but problematic.

The high-stakes field of global risk governance provides some useful clarification into the types of groups in play. Renn (2005) proposes three main groups: experts, stakeholders, and the public. Experts will consist of specialties. Stakeholders are defined as "socially organised groups that are or will be affected by the outcome of the event or the activity" (p. 49). The third main group is the *non-organized* public, which can be split into "the non-organised *affected* public and the non-organised *observing* public" (p. 49). Finally, there are "the media, cultural elites and opinion leaders" (p. 49).

Renn's groupings make the failure rate of conservators sharing treatment decisions (Table 1) even worse than we thought, inasmuch as all the successes listed for sharing with 'museum professionals' do not count in Renn's terms, we were just sharing with our own kind – experts.

Renn (2015) subsequently published an overview of sharing techniques for risk governance decisions, well worth reading for applications to our field (and free online). He structures the consultation process around a hierarchy of three 'challenges': complexity, uncertainty, and ambiguity (the latter covers our issue of value judgements).

Table 3. The three challenges of decision-making for society, the groups that must share them, and the methods used. Abridged from Renn (2015).

	Complexity of the problem	Uncertainty in available knowledge	Ambiguity in social and cultural judgements
	>>>>>> Escalation in group engagement >>>>>>>		
Who shares the challenge	Experts	Experts; stakeholders	Experts; stakeholders; the public
How they address the challenge	"Ask experts for relevant knowledge."	"Involve all affected stakeholders to collectively decide the best way forward."	"Include all actors so as to expose, accept, discuss and resolve differences."
Sequence for the cooperative discourse model, overseen by a team of leaders from each group	Step 2. Experts from multiple disciplines judge each option against each criterion.	Step 1. Ask experts and stakeholders for all concerns and goals; then their criteria for judging options.	Step 3. Randomly-selected citizens evaluate each option (participatory discourse).

He suggests that each successive challenge requires an "escalation" in group engagement (Table 3).

Luckily for us, Renn ends his article with a hybrid model called cooperative discourse, which I think can be scaled to conservation decisions (the last row of Table 3). One dives into the central column – consultation between experts and stakeholders – to establish goals and criteria. Then one goes back to the experts alone who judge the options against all these criteria. Finally, one asks representatives of the public to evaluate the same options in an informal discursive manner. The whole process is guided by a team of leaders drawn from all three groups.

The Delphi method: a secret ballot before sharing

There are many sharing tools and "expert elicitation" tools (Renn, 2015; Kaner, 2014), but I have found the central tactic of the Delphi Method to be particularly powerful, even when used informally. The tactic is the secret ballot. You must collect the judgements of a group, such as scores, weightings, estimates of probability, etc. individually, by some form of secret ballot, before letting them discuss their opinions as a group. These secret votes can then be shared. Individuals with judgements far from the average judgement can choose to explain their vote. Only then should the group seek consensus. This avoids the very common pitfall of group think driven by domineering individuals.

Voting charts

Charting the distribution of individual votes for any numerical judgment helps the group 'see' the degree of divergence or convergence in the estimates. Figure 4 shows the voting distributions for Option 4 in the SCD 2008 case study of Figure 1. Seven people scored each

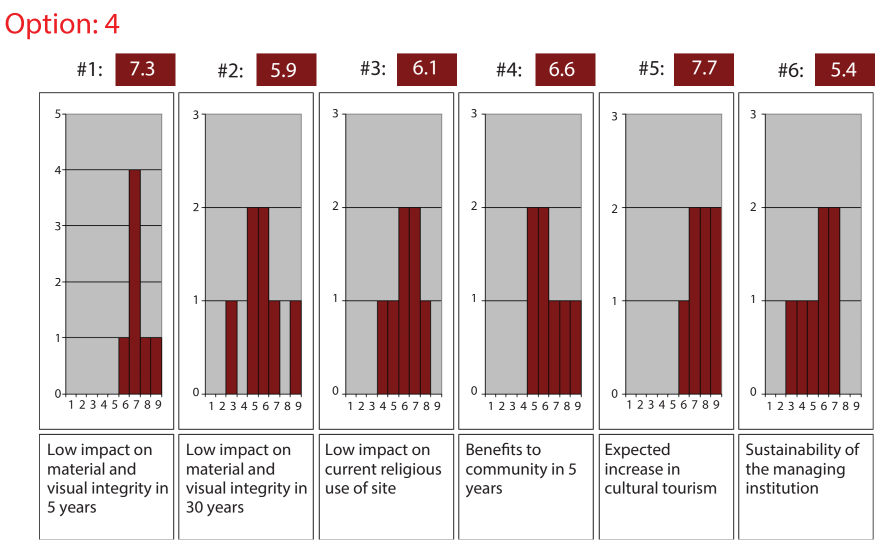

Figure 4. The voting charts for option #4 of the SCD 2008 case study.

option using the six criteria, on a scale of 1 to 9. In Figure 4, one sees that the voting on criterion #1 had the sharpest peak, four people voted a '7', so strong agreement, whereas voting on criteria #2 was twice as widely spread. All the other criteria showed spreads in opinion of 4 to 5 points. The good news is that none of the charts showed votes spreading over all 9 points, and none showed a bimodal distribution (two peaks) which usually signifies a disagreement on what the criterion actually meant. (During this case study, time did not allow iteration of the votes after discussion.)

Participatory decision-making

"Building shared understanding is a struggle, not a platitude" (Kaner, 2014, p. 20). Kaner's book, now in its third edition with a wealth of plaudits, explains the tactics that a facilitator needs to help groups reach sustainable decisions. The primary diagram in his book is a full-page diamond (shaped like one of the blue diamonds in Figure 3). The point on the left represents the beginning of the discussion, the point on the right represents the conclusion. Between the two is a period of divergent thinking followed by a period of convergent thinking. Kaner states that the fundamental mistake is to address difficult decisions the same way as one addresses routine decisions (which Renn (2014) called linear decisions). To find sustainable decisions for difficult problems Kaner insists that we sit in the middle of the diamond for as long as it takes to discover common ground. He calls it the "groan zone". Without shared common ground there will be none of the "insightful collaboration" needed for a sustainable decision. Tactics for difficult decisions, compared to tactics of routine decisions, require a shift from 'either/or' to 'both/and,' from 'analysis of parts' to 'synthesis of a whole.' Sharing must produce long-term unanimity, not just short-term majority rule.

Ideas

Type 1 and 2 thinking

The current model of how we think proposes two kinds of thinking. One is fast, intuitive, confident, and in charge most of the time. The second is slow, deliberative, lazy, and dormant most of the time. Kahneman (2011) has written the most popular book on the subject, and refers to "systems 1 and 2". His text is clearly written, but I find it too skewed towards examples from economics for our purposes. I think Manktelow (2012) provides the more thoughtful perspective for our field. He is especially helpful on the differences between Kahneman, who tends to emphasize the weaknesses of type 1 thinking (its biases) and Gigerenzer, who tends to emphasize the strengths of type 1 thinking (its efficiency).

Heuristics

Most of this section is drawn from Manketlow's (2012) book, where one can find all the many primary sources.

Heuristic does not mean subjective or biased or irrational *per se*. It means a mental strategy for making decisions that is efficient – 'fast and frugal' – and correct most of the time in the context that created it. Three contexts have created three groups of heuristic: evolution of our species, long experience of individuals, and the application of big data techniques.

The first group of heuristics has been learned by our species, and is now hard-wired into our brain's system 1. There are dozens of them, and we use them all the time without effort or awareness. Although they must all have been adaptive from the perspective of the species, many have become 'cognitive biases', flaws from the perspective of the individual trying to be logical. These are organized beautifully in a large graphic under the article "List of cognitive biases" in Wikipedia (2017).

These biases are much studied by economists. Kahneman (2011) developed a model of how we make relative value judgements called "prospect theory". One of its foundations is that we feel losses much more than we feel equivalent gains. In our field, this means that damage to an object (a loss) will weigh more heavily on us and our stakeholders than an equivalent restoration (a gain). This can explain the popularity of 'minimum intervention' since even a small chance of treatment failure seems to outweigh an excellent chance of treatment success. Another foundation is that we judge gains, or losses, relative to what we already possess, or owe. This isn't just the trivial case that $10,000 has more 'value' to us than it does to Bill Gates, but also subtle situations where we spend time and energy to find a store where we can save $10 on groceries but we will not spend the same effort to save $10 on the purchase of furniture, despite the fact that $10 has the same value to us. In our field, if a conservator who is responsible for all the nation's sites is sharing a decision about one site with a community that only possesses that one site, then even if both sides agree on the absolute gain or loss due to some option, the community will feel that gain much more than the conservator, and any loss even more so. When we are sharing conservation decisions, we should be sensitive to phrases such as, "It is the only one we have" or, "I don't want to take any chance of damaging it". Such biases are not errors, they are explanations of legitimate differences in perspective.

The second group of heuristics is the one learned by individuals through long experience (a minimum of 10 years). In the past, this kind of thinking was referred to as tacit knowledge. Experts merge such tacit knowledge with the explicit knowledge of their discipline, even in professions that pride themselves on their objectivity rather than their skills, such as scientists (Collins, 2010). Classic examples in the literature are taken from professions that do pride themselves on tacit knowledge – the fire chief's ability to 'read' a fire and how to attack it; the fine art expert who can 'read' a sculpture as authentic or 'wrong'. Research has clarified that valid heuristics of this type can only emerge for phenomena that actually have a consistent *pattern* that can be *observed*, even if subliminal. The stock market, for example, is *not* such a system. 'Hot' brokers do emerge from time to time, but they are not proof of special pattern recognition, they are equivalent to long strings of the same digit that emerge from time to time in a random number sequence.

What lessons for our shared decisions? I think we need to accept that valid tacit expert knowledge does exist, that it is not subjective in the pejorative sense, but that asking an expert to fully explain how they reached their judgement is of limited use (but worth trying). Scepticism about expertise should be based on two questions: do we think that the phenomenon in question has an observable pattern, and does this person have at least a decade of relevant immersion in this phenomenon. Expert elicitation tools such as the Delphi Method further refine reliability by asking for the opinion of many credible experts, and ensuring that individual opinions are documented before group-think sets in.

The third group of heuristics has been created by researchers who look for patterns in large sets of data. The classic example is a fast three-step decision tree developed to sort cardiac emergency patients into high and low risk groups. This simple decision tree, derived from the analysis of many hospital records, is not only faster and cheaper than traditional and more detailed diagnoses, but also more reliable. Karsten (2016) is developing heuristics for risk assessment of collections. By analyzing many laborious comprehensive risk assessments, she has also found short sequences of simple questions that provide reliable prediction of certain high risks, such as flood damage and fire damage. Sharing during decision-making enlarges the pool of data, and the larger the pool of data, the more likely it is for a valid heuristic to be uncovered which can aid the decision. ("Yes, I've seen that same pattern, I can agree that it's the best indicator we have.")

Reflection, type 3 thinking

"Being intelligent is not the same thing as being smart" (Manktelow, 2012, p. 259).

Two researchers, Evans and Stanovich, propose a revision to the popular two system model, to explain, among other things, the mysterious phenomenon of very clever people believing and doing very stupid things (Manktelow, 2012). Clearly, it would be useful for us also to understand how to avoid such behaviour.

In brief, type 1 thinking is uniform across individuals. It uses heuristics to quickly generate best guesses which are then handed to type 2 thinking which may or may not decide to analyze these guesses further. Type 1 thinking can also send its decisions directly to our beliefs, to what we say, and to what we do. The trick that distinguishes individuals who get beyond these rapid responses is called the 'reflective mind'. Stanovich labels this 'type 3' thinking. Type 3 thinking enjoys being sceptical of type 1's output and asks type 2 to wake up and apply its intelligence to the problem. Type 3 thinking is a trait of an individual's personality; in Manktelow's words, it is the ability to be "open-minded".

The elephants in the room

There is a growing literature called "experimental moral philosophy" (Alfano and Loeb, 2016). Two of its major practitioners have recently published accessible books (Haidt, 2013; Greene, 2013) that I think offer several insights into our topic of sharing conservation decisions. It is not trying to build rules or prescriptions about right and wrong, good and bad, it is trying to understand our moral instincts, our gut feelings, usually by thought experiments. For example, five people are trapped on a railway track, one person is trapped on another track. A train is headed towards the five, but you can pull a switch to redirect it towards the one. What do you do?

For type 1 and 2 thinking about morality, Haidt (2013) has adopted the metaphor of 'the elephant and the rider' within each of us. Our 'elephant' (our type 1 thinking) is fast in providing its 'gut feelings' but it is very difficult to change its opinions, its values. Our 'rider' (the self-aware, type 2 part of our minds) deludes itself that it controls the elephant. Studies show that much of the time the rider is making up a plausible story after the fact, to justify the elephant's choices (confabulation). Scientists mistakenly believe that piling up scientific evidence will convert those who don't believe in climate change. A recent study, i.e. actual evidence, showed that scientific literacy did not predict whether someone in the general population believed in climate change or not. Instead, scientific literacy made opponents on both sides of the debate more certain of their opinion, more passionate (Kahan *et al.*, 2011). In other words, a better informed rider simply becomes a better rationalizer for the elephant and its *a priori* values. When sharing contentious decisions with stakeholders, technical experts, such as conservators, cannot assume that their greater technical knowledge is their most persuasive tool. Worse than that, we must understand that our expertise may be blinding us to the true source of our opposition to the other point of view – our own elephant's values. Is it really facts and reasoning that leads some of us to value original material over original intent in a painting, or deeply buried beliefs?

One method that can help move an elephant's opinions (slowly) is 'framing' the issue in a more agreeable context. For example, let's not speak about that cabal of arrogant museum directors obsessed with blockbuster shows who forced the relaxation of relative humidity standards, and let's begin our discussion instead with the possibility that you can become the hero who gets your museum its environmental certification…now, with that 'in mind', let's reconsider the scientific evidence on the dimensional response of paint!

The sanctity/degradation foundation

Haidt has proposed five modules within the elephant's thinking about right and wrong, which he calls our moral foundations (moralfoundations.org, 2017). I think that one in particular resonates with our field: the sanctity/degradation foundation. I think that all the polemics about 'cleaning controversies' are after-the-fact rationalizing by our rider of the outrage triggered in the sanctity/degradation module of our elephant. This sense of sacrilege is evident in the title chosen for the lengthiest tract ever published in this vein – *The Ravished Image, or, How to ruin masterpieces by restoration* (Walden, 1985).

The "hands-on" blame module

Greene's (2013) specialty is thought experiments, such as the question posed earlier about five people trapped on a train track which you can save by pulling a switch that redirects the train on to a track where one person is trapped. Most people (87 percent) state that they would pull the switch to sacrifice one and save five. Greene calls this the utilitarian decision. But, what if you yourself must push the person on to the tracks to stop the train and save the five. Most people state that they would not push a person to save five others, even though the utilitarian argument is unchanged. When asked this question as well as other dilemmas, medical doctors decide similarly to the general population (don't harm the one). Public health professionals, however, are more likely to make the utilitarian choice (save the five) although they do acknowledge discomfort. I think there are two situations where this public health difference might emerge in our field.

First, conservation professionals have learned to become utilitarian in their judgements since they think of what's best for the long-term greater good (the interests of future generations). It is not unusual for a proposal to have a short-term disadvantage that achieves a long-term advantage. Stakeholders tend to focus on the short-term – their own generation – so they object. Second, preventive conservation (and risk management) are justified as the efficient protection of entire collections instead of the traditional one-special-object-at-a-time perspective of the public (and the bench conservator). Greene wonders whether one learns these professional utilitarian perspectives on the job, or one has them already, and is drawn to the professions that exercise them. Either way, sharing conservation decisions will involve resolving these opposing perspectives.

Finally, by exploring many variations of the 'people on a train track dilemma', Greene has uncovered some of the building blocks of our moral judgements, and it is not good news for our profession. Our 'do-no-harm alarm' is triggered only if the causal relation is simple and direct. Side effects from sending the train down another track does not trigger it. The decision is handed over to type 2, utilitarian thinking – which has no difficulty deciding that one death is better than five. However, the thought of using our hands to push the one person definitely triggers the 'do-no-harm alarm'. Killing the five by doing nothing is too indirect to trigger the alarm. Hence the odd indifference to 'collateral damage'. A conservation treatment is literally the placing of the conservator's hands on a special thing, so the conservator is obviously the cause of whatever sacrilege or degradation occurs. I suspect that the life and death alarm bells that Greene has uncovered can be applied to judgements about things that are 'priceless' or 'irreplaceable' or sacred. If all goes well, we are heroes, if not, we are villains. One benefit of the sharing of treatment tasks is the shared ownership of the results.

Conclusion

When sharing a decision becomes difficult

Research on human reasoning and moral judgements summarized in this article has uncovered a complex but universal set of mental mechanisms that have evolved over millennia, sometimes labelled 'type 1' thinking or more colloquially as our 'elephant'. The research also finds profound variations in the settings of these mechanisms between individuals and between cultures. We can expect, therefore, that if sharing a decision with stakeholders has become contentious, it is probably because of a variation in type 1 thinking between individuals or between cultures. It is important for leaders in the sharing process to understand that judgements based on values or feelings, especially when vociferous, are not something that a person can explain, they can only express. A decision matrix can help the sharing of difficult decisions in two ways: it partitions complex contentious issues into their fundamental value judgements (the criteria) and it captures the strength of each participant's connection to those criteria.

Empathy

Kaner (2014) stresses empathy (putting ourselves in another's shoes) as an essential tactic for participatory decision-making. Technocrats might dismiss this as a touchy-feely platitude, but I think Haidt's work makes it clear that, for value-based decisions, we have no choice but to try and understand our own elephants and those of others.

Accountability

When difficult heritage treatment decisions are being shared, it is usually the case that some, if not all, the actors have significant legal and fiduciary responsibilities. In government, decision tools are invoked for transparency and accountability. One might as well make the best of them rather than consider them a hindrance.

Technical overreach

Technical overreach refers to the tendency of experts to presume control of the whole decision process, to presume to represent the groups affected, such as stakeholders and the public. At best this is naïve, at worst it is offensive.

As Renn (2004) makes clear for the field of public risk decisions, and Chirikure and colleagues (2010) make clear for world heritage sites, it is essential that technical experts bring as much relevant knowledge as they can to the analysis stage, but for the final stages of the decision, they must hand it over to the affected groups. I suspect that our profession is even more prone to overreach because technical issues often merge with value issues within our own domain.

Reflection, the key?

I think that it is obvious that reflective thinking is the key to good decisions. We can recognize it in every culture's aphorisms about wisdom and thoughtfulness, but we can also recognize it in the evidence and advice I have compiled here – it is Kaner's "groan zone" and it is precisely what is missing in Nutt's number one blunder – "rush to judgement". It is, presumably, what was missing in the sharing and decision failures documented in our own field of heritage. After all, we do not think that these failures were due to a lack of intelligence, or evil intentions, do we? If reflection is the key, then a primary purpose of our tools and our tactics must be the facilitation of reflection.

Sharing with several people will always favour reflection: first, there will be a higher chance that someone is innately reflective, and second, there will be a higher chance of initial disagreement, which might then trigger constructive reflection.

Tools, such as the decision matrix, decision tree, secret ballots, and voting charts facilitate reflection by capturing the easily neglected insights of introverts, and allowing complex structures to emerge that belong to the whole group. Software versions of these tools projected on to a shared screen allow a 'fast and frugal' response when someone asks, "What if we change that score?" What-if games are, in fact, the arithmetic version of reflection. They can also maintain a record of every iteration of the reflection process, i.e. don't throw away the drafts.

The good news about reflective thinking is that, unlike raw intelligence, it can be developed beyond our individual limitations, through tools, the questions of others, self-awareness, intellectual humility, and, above all, through the honest sharing of our difficult decisions.

References

Alfano, M. & Loeb, D. 2016. Experimental Moral Philosophy. In E.N. Zalta, ed. *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Caple, C. 2000. *Conservation Skills: Judgement, Method and Decision Making*. London, Routledge.

Chirikure, S., Manyanga, M., Ndoro, W. & Pwiti, G. 2010. Unfulfilled promises? Heritage management and community participation at some of Africa's cultural heritage sites. *International Journal of Heritage Studies*, 16(1–2): 30–44.

Chua, H.F., Boland, J.E. & Nisbett, R.E. 2005. Cultural variation in eye movements during scene perception. *PNAS*, 102(35): 12629–12633.

Collins, H. 2010. *Tacit and Explicit Knowledge*. Chicago, University of Chicago Press.

Driver, J. 2014. The History of Utilitarianism. In E.N. Zalta, ed. *The Stanford Encyclopedia of Philosophy*. Winter 2014 edition. Metaphysics Research Lab, Stanford University. (also available at https://plato.stanford.edu/archives/win2014/entries/utilitarianism-history/).

Greene, J. 2013. *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. New York, Penguin Press.

Haidt, J. 2013. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York, Pantheon Books.

Hedley, G. 1990. Long lost relations and new found relativities: issues in the cleaning of paintings. In B.A. Ramsay-Jolicoeur & I.N.M. Wainwright, eds. *Shared responsibility: proceedings of a seminar for curators and conservators*, pp. 8-13. Ottawa, National Gallery of Canada.

Henderson, J. & Nakamoto, T. 2016. Dialogue in conservation decision-making. Studies in Conservation, 61(sup2): 67–78.

Henderson, J. & Waller, R. 2016. Effective preservation decision strategies. *Studies in Conservation*, 61(6): 308–323.

Hwang, C.-L. & Yoon, K. 1981. *Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey*. Berlin and Heidelberg, Springer.

Kahneman, D. 2011. *Thinking, Fast and Slow*. New York, Farrar, Strauss and Giroux.

Kahan, D.M., Wittlin, M., Peters, E., Slovic, P., Ouellette, L.L., Braman, D. & Mandel, G.N. 2011. *The Tragedy of the Risk-Perception Commons: Culture Conflict, Rationality Conflict, and Climate Change*. Rochester, NY, Social Science Research Network. (also available at: https://papers.ssrn.com/abstract=1871503).

Kaner, S. 2014. *Facilitator's Guide to Participatory Decision-Making*. San Francisco, Jossey-Bass.

Karsten, I. 2016. *When disaster mitigation is a priority: Evidence from risk analysis of rare events*. Presentation at the joint Annual Meeting of the AIC and the CAC, Montreal, 2016.

Kepner, C.H. & Tregoe, B.B. 1976. *The rational manager; a systematic approach to problem solving and decision making*. Princeton, Kepner-Tregoe.

Manktelow, K. 2012. *Thinking and reasoning: An introduction to the psychology of reason, judgment and decision making*. New York, Psychology Press.

Michalski, S. 1994. Sharing responsibility for conservation decisions. In W.E. Krumbein, P. Brimblecombe, D.E. Cosgrove & S. Staniforth, eds. *Durability and Change. The Science, Responsibility, and Cost of Sustaining Cultural Heritage*, pp.241–258. Chichester, Wiley.

Michalski, S. 2008. Social discount rate: modelling collection value to future generations, and understanding the difference between short-term and long-term preservation actions. In *ICOM Committee for Conservation, 15th Triennial Conference, New Delhi, 22–26 September 2008*, pp.751–758. New Delhi, Allied Publishers PVT.

Michalski, S. & Rossi-Doria, M. 2011. Using decision diagrams to explore, document, and teach treatment decisions, with an example of their application to a difficult painting consolidation treatment. In ICOM Committee for Conservation, *ICOM-CC, 16th Triennial Conference, Lisbon, 19-23 September 2011*, pp.1–8. Lisbon, Critério.

Mindtools. Undated. *Decision Matrix Analysis: Making a Decision by Weighing Up Different Factors* [online]. [Cited 15 January 2017]. http://www.mindtools.com/pages/article/newTED_03.htm

moralfoundations.org. 2016. *moralfoundations.org* [online]. [Cited 2 March 2017]. http://moralfoundations.org/

Nutt, P. 2002. *Why Decisions Fail: Avoiding the Blunders and Traps That Lead to Debacles*. San Francisco, Berrett-Koehler Publishers.

Renn, O. 2005. *White Paper on risk governance: Towards an integrative approach*. Geneva, International Risk Governance Council (IRGC). (also available at: http://irgc.org/wp-content/uploads/2012/04/IRGC_WP_No_1_Risk_Governance__reprinted_version_3.pdf).

Renn, O. 2008. *Risk governance: coping with uncertainty in a complex world*. London, Earthscan.

Renn, O. 2015. Stakeholder and Public Involvement in Risk Governance. *Int J Disaster Risk Sci*, 6(1): 8–20. (also available at: https://www.researchgate.net/publication/276386716_Stakeholder_and_Public_Involvement_in_Risk_Governance).

Strang, T. 2004. Introduction: choices and decisions. In *Preservation of electronic knowledge: new knowledge and decision-making= La préservation des document éléctroniques: information récente et prise de décisions*, pp.1–4. Ottawa, Department of Canadian Heritage, Canadian Conservation Institute.

Tzeng, G.-H. & Huang, J.-J. 2011. *Multiple Attribute Decision Making: Methods and Applications*. Boca Raton, CRC Press.

Walden, S. 1985. *The ravished image, or, How to ruin masterpieces by restoration*. London, Weidenfeld and Nicolson.

Watson, S. & Waterton, E. 2010. Heritage and community engagement. *International Journal of Heritage Studies*, 16(1–2): 1–3.

Wikipedia. 2016. Evidence-based medicine. In: *Wikipedia* [online]. [Cited 16 January 2017]. https://en.wikipedia.org/w/index.php?title=Evidence-based_medicine&oldid=755351685

Wikipedia, 2017. List of cognitive biases. In: *Wikipedia* [online]. [Cited 2 February 2017]. https://en.wikipedia.org/w/index.php?title=List_of_cognitive_biases&oldid=759928952

Yoon, K.P. & Hwang, C.-L. 1995. *Multiple Attribute Decision Making: An Introduction*. Thousand Oaks, SAGE Publications.