pHealth 2020
B. Blobel et al. (Eds.)

Evaluations of Decision Support Tools Are Preference-Sensitive and Interest-Conflicted: The Case of Deliberation Aids

Jack DOWIE[a][b][1], Mette Kjer KALTOFT[b], Vije Kumar RAJPUT[c]
[a]*London School of Hygiene and Tropical Medicine*
[b]*University of Southern Denmark*
[c]*Stonydelph Health Centre, Tamworth, UK*

Abstract. The questions 'What constitutes a good health care decision?', and, by extension, 'What constitutes good healthcare decision support?' continue to be asked. The most developed answers focus largely, often exclusively, on the quality of the 'deliberation' component as the determinant of the quality of the decision or decision aid. We argue that these answers and resulting aids reflect the preferences of healthcare professionals and aid developers and that these preferences are closely aligned with their interests. Some interests are material, but many professional, institutional, intellectual, methodological, and ethical. Successful promotion of a particular preference-sensitive, interest-conflicted decision aid does not change its ontological nature. Conflicts of interest are therefore universal and of concern only when this ontology is denied and if aids based on alternative interest-based preferences, such as technologies involving numerical analytic calculation, are subjected to discrimination.

Keywords. Decision quality, decision aid, evaluation, deliberation, three-talk-model, preference-sensitive, conflict of interest

Introduction

Can decisions be judged good or not, and, if so, how might this goodness be evaluated? It is hard to envisage a more important question in health and healthcare. It remains so whether or not we substitute 'the best decision', or 'satisfactory decision', for 'a good decision', or simply ask how a decision can be *evaluated*. In all cases what is at issue is the *quality* of a decision. By extension, the same questions apply to evaluating the *quality of decision support*, such as that offered by technologies in the form of 'decision aids', some computer-based and interactive online, some not.

The questions remain the same whether they are asked at the individual or group level, despite the other vast differences that exist in these two contexts. But we confine ourselves here to the individual context, as personified in the clinical decision. In their 2010 reflections on these questions, Elwyn and Miron-Shatz criticized previous answers and made a proposal that became the basis of considerable subsequent elaboration and development [1, 2, 3, 4, 5]. The initial proposal assumed that a clear three stage sequence

---

[1] Corresponding Author. Jack Dowie, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London, UK WC1H 9SH. jack.dowie@lshtm.ac.uk

exists: *a pre-decisional deliberation process*, an *act of decision determination,* and *post-decisional outcomes*. Previous proposals were criticized, for failing to distinguish between the pre-decisional *process* of deliberation and the subsequent *act* of decision determination. This act was defined as 'choosing an option'.

Surprisingly, and contrary to the expectation set up by the stressing of this distinction, their proposal involved evaluating only the first stage, the pre-decisional deliberative process. They propose

"… a framework for evaluating the deliberation phase which we hypothesize, if done well, *may lead to a good decision determination* … Evaluating the deliberation process, we propose, should comprise of a subjective sufficiency of knowledge, as well as emotional processing and affective forecasting of the alternatives. *This should form the basis for a good act of determination*." [1] (p140).

The proposal leaves the *act of determination* unconsidered - and the question 'Can an *act of determination* be judged good or not, and if so, how might this goodness be evaluated?' unaddressed. This question is not raised in follow-up research either, and work has continued to focus exclusively on establishing the criteria for a good deliberative process. Despite the cautious wording in the above quotations ('may lead to', 'form the basis for') the *quality of the act of determination* is effectively left equated with the *quality of the preceding process of deliberation*.

A revised model was introduced in 2017 with the specific aim of improving the quality of *shared* decision making.

"A new three-talk model of shared decision making is proposed, based on 'team talk', 'option talk', and 'decision talk', to depict a process of collaboration and deliberation. Team talk places emphasis on the need to provide support to patients when they are made aware of choices, and to elicit their goals as a means of guiding decision making processes. Option talk refers to the task of comparing alternatives, using risk communication principles [in discussing harms and benefits]. Decision talk refers to the task of arriving at decisions that reflect the informed preferences of patients. The overarching goal [is] a model that can act as … a primer, a reminder, and a guide to accomplishing shared decision making. Shared decision making is sometimes supported by patient decision aids, which often include risk communication formats. These would fit naturally in the 'option talk' step." [3] (Abstract, Appendix Table A).

Unfortunately, it is left unclear as to how it is possible for the 'option talk' to 'compare alternatives' in a personalized way without introducing the preferences of *this patient* in *this second phase*. Nor how this non-personalized 'comparative talk' can transform the *preferences* of *this* patient from 'uninformed' ones to the 'informed' ones in the 'decision talk' and thence embedded in the decision. Aids to support decision making are mentioned, but they focus on information provision, leaving the crucial task of integrating evidence and preferences unaddressed. So, while the last phase - 'decision talk' - is now more complex than a simple 'act of determination', it is still the case that 'choice between options' just happens. A mind is made up.

1. Preferences and Interests

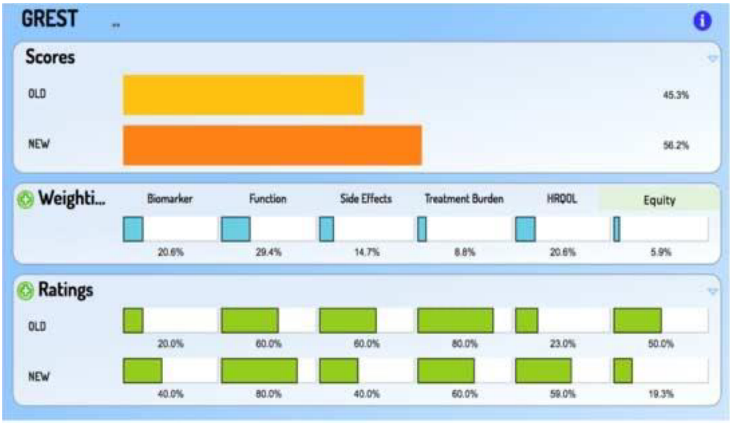
The main objective of this paper is not to criticize the Elwyn models. It is to make clear that, in pointing out what we see as weaknesses in them, we are reflecting our *preferences*. The models are not *rejected* as possible constructions of a decision support and making process, or as possible bases on which the quality of these processes can be evaluated.

But the implicit equation of the *quality of the act of determination* with the *quality of the preceding process of deliberation* is far removed from our preferences. We regard equating 'a good deliberative process' with 'a good decision' - and 'good support for a deliberative process' with 'good decision support' - as reflecting particular 'tastes' in the Bourdieusian sense [6]. While this equating may be a widespread preference, it is not ours and its popularity does not affect its ontology. One important consequence of our very different preferences is that technological support in the form of analytical decision aids becomes virtually a necessity to achieve higher quality, preference-sensitive decisions.

Quality is a *formative* construct. In contrast to a *reflective* construct (such as an 'inflamed appendix'), it does not exist until it is formed, i.e. defined and, in most cases, measured [7]. All serious evaluative measures in healthcare - including evaluations of any product, process or activity - are multi-criterial. They take the form of *indexes* produced by weighting a selected set of criterion *scales*. Using equal weights for the item scales in an instrument is still assigning weights. Note also that the output from a 'profile' instrument - set of unaggregated criterion scales - does not constitute a measure of quality. The selection of the criteria for an index, and the weights assigned to them, reflect underlying *value judgments*, so that all evaluation measures, including decision quality, are both preference-based and preference-*sensitive*.

These preferences - and the concepts and measures into which they are built - are reflections of 'taken-for-granted' assumptions within the field's *doxa.* Many of these assumptions are aligned with *interests* of various sorts. Some may be directly financial, but these can be less important than other material and non-material interests, such as the advancement of a professional career or institutional reputation, or implementing particular political, ideological, methodological or ethical positions (e.g. deontological or consequentialist). When the embodiment of interest-based preferences in quality measures and other evaluations is unconscious, the danger is that this will lead developers and users *to reject (or accept)* alternative measures, rather than acknowledging that they are based on different preferences and interests.

It will make little sense to the reader if the existence of a decision aid based on our very different preferences and interests is not referenced. As has been already suggested, these preferences are for decision support based on numerical analytic calculation, as opposed to the verbal deliberative reasoning format of most current decision aids [8]. Crucially, these preferences involve the provision of an opinion on the decision, explicitly omitted by aids that can, nevertheless, meet the preference-sensitive International Patient Decision Aid Standards checklist [9]. Figure 1 provides an illustration of a generic Multi-Criteria Decision Analysis-Based Decision Support Tool (DST) for use in clinical and commissioning decisions [10]. It can be engaged with at https://ale.rsyd.dk (enter 1513 as survey ID).



Scores	
OLD	45.3%
NEW	56.2%

	Biomarker	Function	Side Effects	Treatment Burden	HRQOL	Equity
Weighti...	20.6%	29.4%	14.7%	8.8%	20.6%	5.9%
Ratings OLD	20.0%	60.0%	60.0%	80.0%	23.0%	50.0%
Ratings NEW	40.0%	80.0%	40.0%	60.0%	59.0%	19.3%

Figure 1: Example of the GREST decision support tool [10]

The nature and scope of 'interests' and 'conflicts of interest', is the subject of a BMJ 'head to head' [11]. Saver [12] and Rodwin [13] represent the opposing positions.

Saver makes the case for widening the legal definition and regulation of non-financial interests to include those which create 'misaligned incentives' problems. And he notes the support for such expansion: '… the editors of PLoS Medicine note that professional affinities and rivalries…scientific or technological competition, religious beliefs, and political or ideological views are often the fuels for (academic) passions and for (research) careers…(and these) interests are perhaps even more potent than financial ones.' [12] (p469).

Rodwin wishes to limit conflicts of interest to those in the current law: "Redefining "conflicts of interest" to include any potential intellectual interests that conflict would make the concept a less practical tool. There is no effective way to eliminate most intellectual conflicts, which are widespread and an inherent part of life. Regulating these potential sources of bias using a conflict of interest framework would burden professionals and institutions for little benefit. On the other hand, financial conflicts of interest generally can be avoided or eliminated." [11] (p2).

Saver dismisses the suggestion that conflicts based on *financial* interests are much easier to distinguish, regulate and eliminate. It ignores the boundary-drawing problem in complex institutional arrangements where the pursuit of non-financial interests can bring financial benefits to both person - and their institution: 'there is no free lunch, and all revenue accruing to anyone or any institution, from any company or governmental program, comes with strings attached.' [12] (p471).

Pointing out that 'value-free evaluations' is an oxymoron is not a recipe for nihilistic subjectivism. Preference-based measures of formative constructs are essential and unavoidable in healthcare. It is entirely normal and to be expected, even encouraged, that different preference-based measures of quality, including decision quality, be proposed, and support for their use sought. The danger lies in failing to acknowledge that that support can actually only be gained on the basis of mutually-shared interest-aligned preferences. There is an almost irresistible temptation to offer a formative measure as being akin to a 'gold standard' [9,14] when its defensible basis lies solely in the intersubjective support of those with similar (enough) preferences and (sufficiently) aligned interests. So, while the danger arises most obviously where there are financial interests underlying the preferences - making them literally value judgments - it also exists, covertly, where the interests are only indirectly material and even if the developers are motivated by the purest of intentions ('pro bono').

2. Discussion

There is a distressingly large amount of work to be done. Two ''crosstalk'' sessions entitled 'What is a 'Good' Medical Decision? Perspectives from Multiple Stakeholders.' were held at the 2015 annual meetings of the Society of Behavioral Medicine Health Decision Making Special Interest Group and the Society for Medical Decision Making. According to the paper reporting on these sessions [15].

"… stakeholders and audience members agreed with the importance of both decision process and outcomes as criteria for a good decision. However, additional issues were identified as being central to answering the question, 'what is a good decision?' These issues included interpersonal factors, structural constraints, affective influences, and values clarification methods. These issues are highly relevant to patients' experiences of both the process and outcomes of medical decision making; yet, each issue has gaps in scientific knowledge that limit the ability to develop a comprehensive definition of good decision making… In order to develop a comprehensive definition of decision quality and to advance the theory and practice of medical decision making, it is critical to consider novel factors that not only shape the process by which patients make medical decisions and the outcomes of those decisions, but also those factors that could serve as new metrics for identifying a good decision. This requires an acknowledgement and better scientific understanding of the real ways that patients make decisions." [15] (pp55, 64).

However, the vast majority of healthcare decisions are multi-criterial, and therefore preference-sensitive, so 'better scientific understanding' can never answer the question. Seeking a preference-free definition of 'a good medical decision' is hunting the mythical Snark of Lewis Carroll [16]. As is trying to establish *scientifically* 'the *real* way patients make decisions'.

We focus on deliberation because it is taken-for-granted within the healthcare professional doxa that it is the only game in town as far as decision making and support is concerned. Regrettably, deliberation has never been subjected to *comparative empirical evaluation* in the way that the scientific orthodoxy would demand for any other process or product (e.g. drug, device) and as is routinely demanded of analytical decision aids. It is therefore particularly important to expose and counter those who benefit from presenting evaluations as truth claims, as is too often the case with so-called 'scientific' assessments. All clashes in *evaluation* - in the placing of value on - are clashes of *preference*. Even clinical trials are preference-sensitive and interest-biased, insofar as they claim to be evaluations of options [17].

3. Conclusion

It is acceptable to define and measure the quality of a decision or decision support tool by the quality of its deliberation component, but this is subject to it being advanced as one of the many possible preference-based definitions and measures. Subject also to not dismissing rival ways of evaluating decision support technologies, except as a declared matter of preference. In two future papers [18,19] we show how the dominant normative and empirical methods of decision aid evaluation reflect undeclared interest-conflicted preferences.

References

[1] Elwyn G, Miron-Shatz T. Deliberation before determination: the definition and evaluation of good decision making. Health Expect. 2010; 13,2: 139–47.

[2] Elwyn G, Frosch D, Volandes AE, Edwards A, Montori VM. Investing in deliberation: a definition and classification of decision support interventions for people facing difficult health decisions. Med Decis Making. 2010; 30,6: 701–11.

[3] Elwyn G, Durand M-A, Song J, Aarts J, Barr PJ, Berger Z, et al. A Three-Talk Model for Shared Decision Making: Multistage Consultation Process. BMJ. 2017; 359: j4891.

[4] Sivell S, Marsh W, Edwards A, Manstead ASR, Clements A, Elwyn G. Theory-based design and field-testing of an intervention to support women choosing surgery for Breast Cancer: BresDex. Patient Educ Couns. 2012; 86,2: 179–88.

[5] Witt J, Elwyn G, Wood F, Rogers MT, Menon U, Brain K/ Adapting the Coping in Deliberation (CODE) framework: A multi-method approach in the context of familial Ovarian Cancer risk management. Patient Educ Couns. 2014; 97,2: 200–210.

[6] Bourdieu P. Distinction. London: Routledge; 1984.

[7] Kaltoft MK, Nielsen JB, Dowie J. Formative preference-sensitive measures are needed in person-centred healthcare at both clinical and policy levels. Eur J Person Cent Healthc. 2017; 5,4: 495-500.

[8] Dowie J. shared decision making is a preference-sensitive formative construct : the implications, Eur J Person Cent Healthc. 2019; 7,3: 506-17.

[9] Elwyn G, O'Connor AM, Bennett C, Newcombe RG, Politi M, Durand M-A, et al. Assessing the quality of decision support technologies using the International Patient Decision Aid Standards Instrument (IPDASi), PLoS ONE. 2009; 4,3: e4705

[10] Dowie J, Rajput VK, Kaltoft MK. A Generic Rapid Evaluation Support Tool (GREST) for clinical and commissioning decisions, Stud Health Technol Inform. 2019; 266: 576–80.

[11] Wiersma M, Kerridge I, Lipworth W, Rodwin M, Should we try to manage non-financial interests? BMJ. 2018; 361: k1240.

[12] Saver RS. Is it really all about the money? Reconsidering non-financial interests in medical research, J Law Med Ethics. 2012; 40,3: 467–81.

[13] Rodwin MA. Attempts to redefine conflicts of interest, Account Res. 2018; 25,2: 67–78.

[14] Dowie J, Kaltoft MK, Why a Global PROMIS® can't be kept. Stud Health Technol Inform. 2019; 262:114–17.

[15] Hamilton JG, Lillie SE, Alden DL, Scherer L, Oser M, Rini C, et al. What is a Good Medical Decision? A research agenda guided by perspectives from multiple stakeholders J Behav Med. 2017; 40,1: 52–68.

[16] Wears RL. The Hunting of the Snark. Ann Emerg Med. 2011; 58,5: 465–67

[17] Dowie J. Clinical Trials and Economic Evaluations? No, there are only evaluations. Health Econ. 1997; 6,1: 87–8

[18] Dowie J, Rajput VK, Kaltoft MK. Measures of decision aid quality are preference-sensitive and interest-conflicted 1: normative measures, Stud Health Technol Inform. 2020?: submitted to e-stc

[19] Kaltoft MK, Dowie J, Rajput VK. Measures of decision aid quality are preference-sensitive and interest-conflicted 2: empirical measures, Stud Health Technol Inform. 2020?: submitted to e-stc