

TOWARD A CONSENSUS ON DECISION AIDING LANGUAGE AND CONCEPTS

Rex Brown, School of Public Policy, George Mason University.

Abstract

Paper discusses Ron Howard's comments and suggestions for standardizing decision aiding terms. In particular, it addresses defects in common practice from the perspective of convenience and accuracy in communication between aiders and decision makers, and proposes solutions.

As a young consultant, I once supported an investment recommendation to a manufacturer with a written decision analysis. The staff group who had hired me seemed impressed, but when they presented it to the president, he snorted "gobbledygook!" and stalked off. This embarrassment might have been avoided if we had used a language he understood. Ron Howard (H) has taken a major step towards developing such a language.

He has pointed out defects in many common terms (such as *expectation*, *dependence stochastic* and *confidence*). He has also proposed a number of useful new terms (e.g. *prospect*, *possibility*, *degree*, *measure*, *clairvoyance* and *present equivalent*). However, some of his substitutions seem uncalled-for (such as *distinction* for *attribute* and *u-value* for *utility*). For others there seem to be better alternatives available (e.g. *average over e-value* and *covariation over relevance*).

1 Selecting concepts

Clarifying concepts and naming them are distinct issues. Concepts need to be precise, but terms do not. An arbitrary name does not affect the precision of a concept; names are chosen on different grounds.

1.1 Concepts not worth naming?

H's desiderata for concepts to name are unexceptionable and his selection meets them. In essence, concepts should be logically sound, useful to decision aiding discourse, and susceptible to precise definition. However, he goes on to specifically *exclude* other concepts as failing these tests. I agree with some of his exclusions (e.g. *confidence*), but not all (e.g. *uncertainty about probability*—see below).

1.2 Personal probability

In particular, H objects to "adjectival" probability (but I take it he does not have in mind *all* adjectives, as in "X's prior probability"). In the main I agree with him. As a

personalist¹, I agree that the probabilities we work with are *normally* personal in that they represent some individual's uncertainty, however well it is documented². However, it is often helpful to distinguish personal from other types of probability (when there is ambiguity).

1.2.1 Reference to other paradigms

In decision aiding discourse, we often need to communicate with (or about) *frequentist* and other long-established schools of “probability”³. Christians discussing religion need a term other than “the faith” to distinguish theirs from Buddhism. By the same token, I favor adding *personal* to *decision analysis* (PDA); to avoid the presumption that what H calls *decision analysis* is the only proper way to analyze a decision.

Sometimes PDA can make good use of work by non-believers⁴. Millions of dollars are spent annually on “Probabilistic Risk Assessments” to support nuclear safety decisions. For institutional and (misguided) technical reasons, these assessments are essentially frequentist and attempt to *exclude* personal judgments. No individuals need to adopt the “probability estimates” used (and usually are wise not to), so they are not personal probabilities. Nevertheless, PRAs often contain useful material, such as experimental results that, if properly interpreted can be adapted for PDA.

1.2.2 Personalist “Impersonal” probability

More controversially, I believe, there is a useful interpretation of *impersonal* probability quite consistent with the personalist paradigm. I have observed that almost everyone who talks about risk—including personalists “who should know better”—cannot resist referring to “*the*” probability of a unique hazard, as if it had an existence independent of any one person's belief. (Listen to your colleagues—you may be surprised!). Is it possible that we are all *not* misguided?

I have attempted elsewhere⁵ to define the *ideal* probability of a possibility, as a consensus among rational judges based on *all* feasibly achievable knowledge—say the result of an unlimited research budget. (Not “perfect information”, which would remove all

¹ In the Savage (1954) tradition, typified by the “maximization of subjective expected utility”.

² For example, “heads” in a coin toss, which is why I favor *personal* over *subjective*, which suggests undocumented judgment. Even if a probability is developed for an *institutional choice*, there will normally be some individual, say the decider or recommender, who adopts the probability as his own.

³ At a 1959 meeting of the Royal Statistical Society George Barnard, a leading classical statistician, took me to task for using the word *probability* for a statement of personal belief. “That's not a real probability” he said. “Fine, then let's call it “credence”, and for a while I did.

⁴ Non-personalist decision analysis may actually be reasonable, when contentious public decisions are subject to outside scrutiny. For political or legal reasons, choices must be seen to be untainted by the “judgment” of a possibly interested party. Decider judgment on some issues may be replaced by expert panels, selected and elicited by a reviewable procedure. The individual probabilities may be personal, but their aggregation (e.g. Delphi) will not be, nor will the overall decision analysis.

⁵ Brown 1993

uncertainty and produce a probability of one). It may not exist in all cases, and in any case, the concept needs testing and development⁶.

However, as a consultant, I have often found it an invaluable construct in high-level nuclear risk management. For example it was used to help a Regional Administrator at NRC to decide whether to close down a reactor or to wait for the results of a safety study⁷. Ideal probability served as a tighter, and therefore more useful, bound than perfect information on the value of research.

1.2.3 Probability as just a mathematical construct

A probability statement can have two quite distinct aspects. It has an interpretation (such as personal), and it obeys a formal calculus as presented in standard textbooks⁸. It is unnecessary to burden the name *probability* with both cognitive and mathematical properties.

I suggest that adherence to mathematical rules be the defining property of *probability*. I.e. that the term probability be interpretation-free, until qualified as personal, frequentist, etc. (The qualifier can be dispensed with if the interpretation is clear from context.) “Probability theory” refers only to mathematical rules, and has nothing to do with the interpretation of the arguments. Thus Bayes’ theorem is *probability* theory, of which one application is to update *personal probabilities*.

1.3 **Uncertainty about probability**

I accept H’s clairvoyancy test, viz. a probability can only apply to a possibility that a clairvoyant could resolve. However, I hold that a personal probability passes that test since the possibility is some judge’s assessment, i.e. a *psychological event*. I can thus be uncertain, say, about what X’s probability is now, or about what my own probability will be after research. The clairvoyant can tell me what it is (even though he “knows nothing about probability”). Uncertainty about ideal probability, discussed above, is comparable.

2 Naming concepts

2.1 **Contentious current terminology**

I agree with most, but not all, of Howard’s objections to common terminology, illustrated in the opening section of the paper.

⁶ A clairvoyant can tell us whether there *exists* an adequately unique ideal impersonal probability in a given case (as well as what the probability is). Thus the construct may not always be applicable. An impersonal ideal probability would *not* be uniquely defined if all accessible evidence were still too little to swamp personal priors (say, for the probability of life in the galaxy). Adequate consensus might be the mean of interpersonal probabilities, if these were tight and stable. It could also shift with the information technology (say, if a galactic life detector appeared).

⁷ Brown and Ulvila 1994

⁸ Kendall & S 1961.

2.1.1 *Utility*

I don't share H's objection to the word *utility*—and certainly wouldn't replace it with *u-value*, which I find quite obscure. Utility has been used for a broad class of measures of “satisfaction”, both ordinal and cardinal. This includes our personalist utility, where satisfaction incorporates risk attitude, operationalized in terms of a reference gamble⁹. When necessary (e.g. when talking to economists), we can distinguish it from, say, “value functions”¹⁰ by a qualifier, such as *gamble utility*.¹¹

I disagree that “in a deterministic world, a u-curve (gamble utility) is useless”, since it is the special case of an “uncertain” world where probabilities are all one. Gamble utility is perfectly appropriate to assigning preference among any options—whether their prospects are certain or uncertain.

2.1.2 Alternatives to H's replacement terms

I concur with many of H's replacements for objectionable words (such as *probabilistic* for *stochastic*). However several of them can be significantly improved upon.

2.2 ***Appropriate connotation?***

Several of H's proposed terms appear to be obscure or misleading for the intended audience. The most serious misnomers H identifies are those that connote a specific *wrong* concept.

2.2.1 *Relevance*

Relevant has too broad an everyday meaning (as in “relevant argument”).¹ I suggest covarying and covariation. Interdependent might also do, preserving some continuity with tradition. Either case reflects appropriate symmetry, which dependence lacks. *Dependence* could be retained as causal dependence.

2.2.2 *E-value*

E-value (like *u-value*) is obscure and without any ready connotation to guide understanding (as well as possible confusion with e-commerce, etc.). I suggest *average* as an alternative.

2.2.3 *Distinction*

Distinction normally connotes difference between two or more items, as H well illustrates in his abstract: “We propose replacing ... any *distinction* between risk and uncertainty”. I prefer the familiar *attribute*, but without limiting it to *criteria*, as it has usually been used in Multiattribute Utility Analysis. Unlike a distinction, an *attribute* can have the same value for several objects.

⁹ General utility may be ordinal, Gamble utility is cardinal, but ranks certain prospects the same as other utility measures. No utility has intrinsic meaning; scales are defined by arbitrary anchor points.

¹⁰ As used by Keeney and Raiffa, 1976.

¹¹ Perhaps a new term can be fashioned from, say, a counterpart foreign term.

2.2.4 Ambiguity

Sometimes, H attaches a narrower meaning to a familiar term than it connotes in general use, e.g. *probability*, *rule* and *decision analysis*. An adjective in cases of ambiguity may solve the problem. Thus, *personal decision analysis* need only be used once in an exposition, and then give way to *decision analysis* if it is clear there is no question of any other type of decision analysis.

Sometimes, an H term may need its scope clarified. I would specify that *prospect* is a consequence measured from what Schlaifer calls a *fixed base* (e.g. money in the bank). If the consequence were measured from a *floating base*, i.e. the prospect resulting from a particular option such “do nothing” (e.g. how much money a certain deal will make). I would call that an *impact*.

2.3 Other Desiderata

Two less critical desiderata can be added to “appropriate connotation”: conformity to established practice and economy of expression

2.3.1 Established practice

However intrinsically sound a terminology is, its adoption will be impeded if it runs counter to well-established practice, particularly if that practice extends beyond the audience reformers can reach and influence. A switch in terms must be important enough to justify overcoming resistance (as I believe is the case with *expectation* and *dependence*). In any case switches would certainly need influential cooperation, especially within the statistics community

If the replacement term already has a conflicting meaning, resistance will be particularly great, since a replacement for the old meaning must be coined. *Likelihood* would do fine for *personal probability*¹², but it has been pre-empted by statisticians for the attribute I would call *diagnosticity*. Objectivist *confidence interval* would be perfect for what we have had to call *credible interval*¹³: it connotes personal judgment. In fact *confidence interval* is already widely misinterpreted as a credible interval (sometimes disastrously as H illustrates)¹³.

2.3.2 Economy of expression

Replacing a cumbersome two words by one is a modest but worthwhile improvement. E.g. replacing *perfect information*. by *clairvoyance* and *decision maker* by *decider*¹⁴.

¹² In fact it is enshrined in nuclear regulation in a probabilistic sense. 40CFR191 reads in part “the likelihood of radioactive emissions (limit) must not be greater than exceeding 10%”.

¹³ . It would also be convenient if we could replace “equivalent to” by “indifferent to”, for similar reasons, but the improvement would not be enough to overcome conflicting everyday grammar.

¹⁴ Which French decision aiders have done with “decideur”.

3 Discussion

3.1 *Developing a definitive glossary*

3.1.1 Prescriptive language

What is needed is a language that real people can use and find useful¹⁵. It is a question of psychology as well as logic, like any other decision aiding tool.

3.1.2 Audience effect

Miscommunication between decision aider and the real decider (rather than the agent-client) remains a prime cause of failed aid. Solving this problem may be the most valuable contribution of a decision aiding language that is accessible to the layman. The *decider-client* is often a busy top executive, often capable of grasping the ideas, but impatient with esoteric language. The *agent-client* may be a technically oriented buying service (e.g. in-house management science group) for the decider-client. Agent staff may be prepared, even enthusiastic, to learn our jargon, but busy deciders certainly aren't.

My own decision aiding experience has made me particularly sensitive to this issue, since I have been working largely with the executives, often quite senior, who make the decisions, rather than their agents. I am quite sure I would lose (in one sense or the other) decider-clients if I used terms like *u-value* or *distinction* on them. My teaching has also been to executive oriented students, such as MBAs and MPPs, rather than, say, engineers or trainee decision aiders.

I agree that the glossary for professionals should generally be the same as for laymen (with additional specialized vocabulary), though some case might be made for adapting language to audience. Paradoxically, it may be easier for lay executives and executives-in-training (e.g. MBAs) who have not been conditioned by past language to adopt new terms, than statistical sophisticates.

3.1.3 Expanding the glossary

H's concepts are not (and do not purport to be) exhaustive. They do not, for example, address conflicting criteria, rating scales, or decision aiding methods he does not favor¹⁶, or concepts that he does not consider "fundamental"¹⁷. This does not matter. They are useful first steps in a journey of 1000 miles. Other authors can fill the gaps.

Appendix A attempts build on his efforts by addressing some of the gaps. It proposes an expanded but still provisional glossary of terms on which there is not yet professional consensus. It gives an indication of which of H's terms I would adopt and which I would replace, as well as concepts H has *not* addressed that I have found important in decision-aiding communication.

¹⁵ Annals? Acta?

¹⁶ Such as Recognition Primed Decision [Klein or Oresanu]

¹⁷ Though as Amos Tversky has said "You don't have to finish the foundations before you start on the roof (private communication 1992).

3.1.4 Validating glossary empirically

In keeping with the useful principle of *plural evaluation*, I favor using both internal and external validation in discriminating among terms. The discourse between H and myself has been conducted largely at the level of *a priori* argument¹⁸. There is an alternative approach to testing the suitability of terms and concepts by empirical validation, i.e. observing performance in practice. (What we English infelicitously call “suck it and see”). It may well be that, as H surmises, the “u” in *u-value* will enable people to “make the connection (with utility curve)”. External validation by experiment or other empirical observation could confirm or refute this.

3.2 Codifying the glossary

3.2.1 Consensus

At some stage it would be valuable to have an “authorized” glossary that newcomers to the field and others could be taught. Initially, it might focus on personal decision analysis methodology, but eventually it would need to be broadened to include other decision aiding approaches.

My own forthcoming decision analysis textbook¹⁹ will propose such a glossary, but I would like to make it conform to some consensus, so early feedback from colleagues would be much appreciated.

Technical vocabulary in a developing field like decision aiding is a notorious source of contention. Perhaps some “Committee on Terminology”, set up by the Decision Analysis Society and possibly other influential parties, could approve and publish a glossary, to give it a seal of approval

3.2.2 Adoption

The main stumbling block in the way of enhancing language *practice* may not be establishing a sound language, but achieving its widespread adoption. The hoped-for universal general language Esperanto, developed in the 1940s was agreed to be a vast improvement over, say, English. However, it has sunk with barely a trace, because people clung to the familiar, which they could get by with.

The same could happen with those parts of a decision aiding language that have become embedded in professional practice (such as *expectation*), however compelling the counter-arguments. It is critical that a new glossary be adopted by the main sources of new decision-aiders, notably the Stanford, Duke and Harvard, and by major authorities in the field, such as Howard Raiffa, Ralph Keeney, Bob Clemen and Ron Howard himself.

Controversy may also be resolved by “natural selection”. Whatever high priests of the profession may decree, the final arbiter of accepted practice will be usage. We may propose vocabulary, but the god of the market will dispose. We will choose among

¹⁸ Brown 1992; Howard 1992.

¹⁹ “Tools of Rational Choice” Wiley, in preparation.

“subjective expected utility”, “e-value” or “personal average utility” by voting with our typing fingers...

Acknowledgements

[Pratt, Watson, Ulvila, Chinnis].

ABOUT THE AUTHOR

The author has alternated careers in academia (including faculties of business, statistics, psychology, engineering and public policy) and decision aiding consulting. His clients have included presidents of Ford and Firestone, heads of reactor regulation at NRC and nuclear waste management at DOE, and assistant secretary of defense.

REFERENCES

- Brown, R.V. The state of the art of decision analysis: A personal perspective. *Interfaces*, **22**, 6 Nov-Dec 1992, 5-14.
- Brown, R.V. Impersonal Probability as an Ideal Assessment based on Accessible Evidence: A Viable Construct? *Journal of Risk and Uncertainty*, *7*:215-235, 1993.
- Brown, R.V. Assessment uncertainty technology for making and defending risky decisions. *Journal of Behavioral Decision Making*, *3*, 213-228, 1990b.
- Howard [this issue of JDA]
- Howard RA. Heathens, heretics and cults: the religious spectrum of decision analysis. *Interfaces*, **22**, 6 Nov-Dec 1992
- Keeney, R.L., and Raiffa, H. *Decisions with multiple objectives, preferences and value tradeoffs*. New York: Wiley, 1976
- Kendall, M.G. and Stuart, A. *The advanced theory of Statistics*. London: Hafner, 1961.
- Lindley, D.V., Tversky, A., and Brown, R.V. On the reconciliation of probability assessments. *Journal of the Royal Statistical Society, Series A*, 1979, *142*(2), 146-180.
- Savage, L.J. *The foundations of statistics*. NY: John Wiley and Sons, 1954.