Beyond comprehension: Figuring out whether decision aids improve people's decisions

A married couple in their mid-40's with two young children ask their financial advisor whether they should increase the percent of their assets placed into high yield stocks. A seventeen year old high-school student meets with her guidance counselor for advice on where to apply to college. A sixty five year old man with a non-metastatic prostate cancer asks his physician whether he should have his prostate surgically removed.

Each of these people is facing what is known as a "preference sensitive decision," where the right choice depends in part on that person's specific preferences (O'Connor et al., 1999). The best investment choice, for instance, depends on a given person's risk tolerance; the best college choice depends on a person's preferences for big cities vs. small towns, liberal arts vs. engineering classes; and the best approach to prostate cancer will depend on how concerned a man is about the risk of impotence or incontinence from treatment.

In each of these cases, the person making the decision is looking for help, from a neutral party—someone who can help them make a decision that is consistent with their underlying goals and preferences. Which raises an important question: how do decision counselors know when they have improved people's decisions?

This question is important because, when left to their own devices, people won't always make the right decision. That is one of the reasons why people seek advice from lawyers, counselors, physicians and financial advisors—they recognize that they will have a hard time getting adequately informed about the issues relevant to the decision at hand, and therefore turn to experts who can help them. In the face of important,

preference-sensitive decisions, keep the job of a decision counselor should be to help decision makers comprehend information about their alternatives and integrate their information with their individual preferences.

In this article, however, I will show why decision counselors need to go beyond helping people comprehend their decision alternatives. The field of judgment and decision making is replete with examples of people who comprehend their decision alternatives and nevertheless make bad decisions.

In this manuscript, I will discuss these important issues in the context of medical decision making, a setting in which decisions often have unusually high stakes, involve complex choice sets, and typically do not provide the decision makers, patients, with the ability to fully inform themselves about their alternatives. The decision makers, in other words, need help making the decisions. I will specifically focus on a growing movement within medicine to provide patients with decision aids (DAs)—structural educational materials designed to inform people about their decision alternatives. While the context of this paper will focus on medical decision making, the lessons I draw from medical decision making, and the preliminary criteria I develop for judging when a decision has been improved, are relevant in other domains where people face high stakes decisions and need help sorting through complex information.

My goal in this article is to explain why the criteria that shared decision making experts have been using to test DAs in health contexts—whether DAs increase knowledge and reduce decisional conflict—are inadequate for determining whether a given DA actually improves people's healthcare decisions. In critiquing these criteria, I evaluate the strengths and weaknesses of seven additional criteria. I conclude that no single criterion is sufficient for evaluating a DA but, instead, that we need to utilize a broad array of testing standards in order to judge whether a specific DA improves people's decisions.

Healthcare Decision Aids: Structure and Evaluation

Recognizing that physicians are not always effective decision coaches, a movement has grown within healthcare to supplement physician communication with "decision aids" (Bekker, Thornton, Airey, & al., 1999; Molenaar et al., 2000; O'Connor et al., 2003). These DAs are patient educational materials, informed by decision analysis, that structure information in ways that make patients aware of the tradeoffs inherent in their treatment choices – explaining, for instance, the possible outcomes of treatment A vs. treatment B, and the likelihood of each outcome. These DAs also strive to activate patients, showing them the important role that their own preferences should play in determining their treatment choice.

DAs have been rigorously tested in randomized trials, and have been shown to increase patient knowledge and satisfaction with their decision, while reducing decisional conflict (A. M. O'Conner et al., 2004). Indeed, DAs are typically judged as effective or ineffective in large part based on these criteria. These criteria, then, are a good starting point for any assessment of how to help people make good decisions.

1. Knowledge and comprehension:

Healthcare decision aid developers have placed a great emphasis on testing whether DAs adequately inform patients about their healthcare alternatives. Indeed, DAs are evaluated during development for balance, thoroughness, and comprehensibility. Constructing an informative and comprehensible DA is often challenging, forcing DA developers to make difficult judgments about how much information to include, how to help people understand probabilistic outcome data, and how to engage people in the information without overwhelming them or boring them. Often DA developers refine materials through focus groups and cognitive interviews. The best DAs are even pre-

tested by literacy experts, to make sure they are not written above a 7th grade reading level. The end result, typically, is a high quality product that dramatically increases patients' knowledge of their healthcare circumstances and their treatment alternatives.

2. Decision conflict & decision satisfaction:

Healthcare DA developers have also contended that DAs should increase decision satisfaction while reducing decisional conflict (O'Connor et al., 1999). They define decisional conflict as "the uncertainty about which course of action to take when choice among competing actions involves risk, loss, regret, or challenge to personal life values" (A.M. O'Conner, 1993). Signs of decisional conflict include "verbalized uncertainty, expressing concern about undesired outcomes, wavering between choices, delaying decisions, questioning personal values, being preoccupied with the decisions, and feeling emotionally distressed by the decision" (A. M. O'Conner, Jacobsen, & Stacey, 2002). And they have developed a measure of such conflict, which they contend that a good DA will reduce (O'Connor, 1995). Along similar lines, developers of healthcare decision aids have promoted the idea that good DAs will increase patient satisfaction with the decisions they make (O'Connor et al., 1999) and again have developed a scale to measure such satisfaction (Holmes-Rovner et al., 1996).

Inadequacy of these criteria

The shared decision making community in medicine has largely assumed that if you give decision makers freedom and information, they will experience high satisfaction with their decisions, will be unconflicted about their choices, and will make decisions that reflect stable underlying preferences, or values. As readers of this book are undoubtedly

aware, a wealth of studies have demonstrated that these assumptions are often false, and that free and informed decisions are not always good decisions.

1. Knowledge doesn't protect people from bias

The judgment and decision making literature is replete with evidence of biases that can unduly influence even the most informed decisions makers. I illustrate this problem with a study my research team conducted on a DA we designed to help women contemplate whether to take Tamoxifen to prevent breast cancer. Tamoxifen is a hormone-like medication that was initially used in breast cancer patients to reduce the chance that the breast cancer will return. Clinical trials have more recently demonstrated that Tamoxifen can be used in high risk women, to prevent them from developing a *first* breast cancer (Day, 2001)**.** For instance, a woman with a 6% chance of developing a first breast cancer in the next 5 years (based on things like family history and age) can cut that risk in half by taking Tamoxifen. But this medication is not harmless. Women taking Tamoxifen have a chance of developing blood clots, endometrial cancer, hot flashes, and cataracts. In short, the decision to take Tamoxifen to prevent a first breast cancer is by everyone's reckoning a preference-sensitive decision.

In developing our DA, we wanted to test how much women's attitudes toward Tamoxifen would be influenced by subtle changes in how we presented them with information about its risks and benefits. For example, we varied the denominator we used to illustrate the frequency of Tamoxifen's side effects. We informed some women that 17 out of 100 women taking Tamoxifen would experience cataracts, and others that 170 out of 1000 women would experience this side effect. We also varied whether women learned first about the least common side effects of Tamoxifen or the most

common side effects. We found that women's attitudes toward Tamoxifen significantly varied depending on these two subtle manipulations. Women were more concerned about Tamoxifen's side effects when they were told how many out of a thousand women would experience these side effects, and also more worried when the last side effect they learned about occurred with a high-probability. Nevertheless, women's knowledge of Tamoxifen's side effects was not influenced by either the denominator we chose in describing the risks, nor the order in which we presented the risks (Zikmund-Fisher, Fagerlin, Roberts, Derry, & Ubel, 2008).

In a famous study, McNeil and colleagues discovered that people are more willing to undergo a surgical operation with a 90% survival rate than one with a 10% mortality rate (McNeil, Pauker, Sox, & Tversky, 1982). Framing outcomes in terms of survival increased the desirability of the intervention, without altering people's comprehension. Either way of framing the information—90% survival or 10% mortality—will lead to similar comprehension, while causing people to make different decisions, depending on whether the framing triggers people's aversion to losses (Tversky & Kahneman, 1981).

These studies illustrate an important challenge facing DA makers: seemingly neutral manners of presenting information can bias people's judgments and decisions, even if at the same time they increase people's knowledge of their decision alternatives. Knowledge, therefore, is a necessary part of any good decision, but it is not sufficient (Kennedy, 2003). A good DA should not only help people comprehend their choice alternatives, but should also do so in a way that will minimize decisional biases.

2. Is conflict such a bad thing?

As mentioned above, healthcare decision experts also evaluate DAs according to whether they reduce decisional conflict. But their evaluation criterion rests on the assumption that reducing conflict ought to be a goal of DA. This assumption is questionable.

Consider two women deciding on early-stage cancer treatment. The first woman searches the internet, and finds a company website that explains why its pill is the best available treatment for this cancer. She to asks her doctor to prescribe the pill. She feels comfortable about her decision, and has no decisional conflict and high decisional satisfaction. By contrast, the second woman receives access to a DA designed by a non-profit foundation committed to helping people make informed choices. The DA provides information on several treatment options. With the help of this DA she learns about the risks and benefits of the treatments, and picks the choice that she thinks fits her preferences best. But even though she has dismissed alternatives that she realizes would not suit her preferences well, she is still not sure she has made the right decision, and feels conflicted about what the best is.

This second woman has higher decisional conflict and lower satisfaction than the first and, therefore, by the standards that dominate the medical decision making world, would be viewed as having used a *worse* DA. By contrast, the person who relied upon the industry information, which was quite persuasive, felt unconflicted about her decision. Thus, by this criterion, the industry "DA" would be judged superior to the less biased one. In judging DAs, we need to be open to the idea that a good decision may still leave people with substantial decisional conflict (Nelson, Han, Fagerlin, Stefanek, & Ubel, 2007).

Additional criteria for judging "good" decisions

Decision making experts in healthcare have adopted criteria for judging DAs that follow closely upon models of rational choice. The shared decision making community in medicine has largely acted as if a free and informed decision is a good one, and therefore a DA that enables people to comprehend their healthcare alternatives, and thereby reduce decisional conflict, is a good DA. But what other criteria might we consider for judging whether a DA has improved people's decisions? Below I discuss 7 additional criteria (see Table 1), and discuss their role in evaluating whether an intervention like a DA has improved people's decisions.

1. The expected utility criterion

If the goal of a DA is to help people make decisions consistent with their preferences, then a DA could theoretically accomplish this goal by quantifying people's preferences and calculating the expected utility of each decision alternative. If a decision analyst knew the utility that a prostate cancer patient placed on impotence and incontinence, for instance, she could insert those utility values into a decision model and tell that man which treatment maximizes his expected utility.

This approach has the advantage of "doing the math" for people, when faced with decisions that are too complex to otherwise grasp. Indeed, decision analysis has been proposed as a method for determining whether a given decision is preference sensitive (P. A. Ubel & Loewenstein, 1997). If altering utilities across plausible ranges does not alter which alternative has the highest expected utility, then the decision is not preference sensitive.

However, in cases where the utility values do matter, I am concerned that our ability to quantify utility values is often too imprecise to determine which alternative is

best. For example, the most common method of measuring health-related utilities is the standard gamble method, derived from the axioms of von Neumann and Morgenstern (Von Neumann & Morgenstern, 1947). In a standard gamble utility elicitation, a man might be asked what chance of death he would take to rid himself of impotence (Gold, Siegel, Russell, & Weinstein, 1996). A closely related alternative known as the time-tradeoff method would ask that man how many of his remaining years of life he would give up in order to avoid impotence (G. Torrance, 1976; G. W. Torrance, Thomas, & Sackett, 1972). These preference measures are extremely difficult for many people to understand. For instance, people who have difficulty with probabilities and frequencies are often confused by the questions, and thus give non-sensical answers (Woloshin, Schwartz, Moncur, Gabriel, & Tosteson, 2001). People often also raise moral objections to these questions (Baron & Spranca, 1997). In time tradeoff elicitations, it is common for people to say they would not give up any time to improve their health, even when considering horrendous health care problems. They respond this way because they feel it would be wrong to give up any amount of their lives, the moral equivalent of committing suicide.

Utility elicitation measures are also potentially flawed because they are influenced by affective forecasting errors. Most members of the general public, for example, assume that physical disabilities would have a much larger impact on their emotional well being than people with those disabilities report (P. A. Ubel, Loewenstein, Schwarz, & Smith, 2005). Therefore, even people who understand utility elicitations, and who have no moral objections to the questions, may nevertheless bring inaccurate beliefs about the health state question to mind, thereby biasing their responses.

These affective forecasting errors would likely plague an alternative preference measurement method—conjoint analysis. In conjoint analysis, decision makers are given a series of pair-wise choices, with random variation of specific attributes for each choice

(Green & Srinivasan, 1978). Across conjoint choices, it is possible to analyze the weight that any given attribute contributes to a person's decisions. These weights can then be plugged into a decision analysis as utility values. Conjoint analysis avoids some of the problems of standard gamble and time trade-off utility measures. The conjoint approach doesn't usually raise moral objections from participants, and doesn't involve the use of difficult concepts like percentages. But the conjoint approach still opens the door to affective forecasting errors. If a person incorrectly predicts that a colostomy would make him miserable, for example, then the presence of a colostomy in one of the choice pairs will dominate the person's conjoint decisions. Conjoint analysis is also flawed because it still requires people to make complex decisions. A conjoint analysis might ask people to compare two choices over five attributes. This type of decision is susceptible to a range of well-known decisional biases.

Thus, decision analytic models can help determine whether a decision is preference-sensitive, and can even help reveal which preferences are most influential in determining the best choice. But the measurement of preferences—that is, of health-related utilities—is often too imprecise for this criterion to point toward the "correct" choice.

2. Reduction of mispredictions

As the previous discussion illustrated, mispredictions stand as formidable barriers to optimal decision-making. People can mispredict how decision-relevant outcomes will affect their emotional lives (P. A. Ubel, 2006). For instance, when making investment decisions, people may seek out high risks in hopes of transforming their lives with a huge payout. Similarly, when deciding how to treat their inflammatory bowel disease, patients may inappropriately eliminate surgical options out of a mistaken belief that a colostomy would make them miserable (D. M. Smith, Sherriff, Damschroder,

Loewenstein, & Ubel, 2006). People can also mispredict the non-emotional consequences of specific circumstances. For example, people with kidney failure overestimate how much a successful kidney transplant will improve their job prospects (D. Smith et al., 2008). A well-designed DA should reduce or eliminate such forecasting errors. But what would it mean for a DA to do this?

At a minimum, a DA should provide people information about the emotional and non-emotional consequences of specific decision-related outcomes. This might sound trivial, but this approach has not historically been the norm for DA developers. Financial advisors might advise people on the chance of losing half their savings versus the chance of quadrupling their investment. But I expect that few such advisors provide clients with information on how happy they are likely to be with each of these outcomes. I expect most financial advisers don't even recognize the true relationship between net-worth and happiness (Diener & Seligman, 2004). Similarly, in healthcare DAs, patients are presented with neutral language describing specific health-related outcomes. But they are not typically given complete information about the consequences of these outcomes.

DAs could reduce mispredictions by giving people "the answers" – telling them, for instance, how happy people are who have experienced the circumstance in question (Gilbert, Killingsworth, Eyre, & Wilson, 2009). A financial advisor could give a client information about the average happiness of people with a net worth of $500,000 versus $2.5 million. A healthcare DA could report on the average happiness level of people with and without kidney failure.

Rather than give people the answers, DAs could reduce mispredictions by using de-biasing techniques that help people correct affective forecasting errors themselves. For example, Wilson and Gilbert asked people to write out a diary of what their days would be like following a college football loss, and in doing so discovered that people

became more accurate at estimating their moods on those days (Wilson, Meyers, & Gilbert, 2003). Similarly, my colleagues and I asked people to think about how they have responded in the past to emotionally salient circumstances, and this reduced affective forecasting errors for experiencing a severe disability (PA Ubel, Loewenstein, & Jepson, 2005).

Thus, we have two ways to use DAs to reduce mispredictions. The challenge will be to determine whether either of these techniques works. This challenge will not easily be met. Suppose, for instance, we tell decision makers that people with kidney failure are, on average, almost as happy as people with normal kidneys (Riis et al., 2005). They could respond to this information one of several ways:

(1) *Complete disbelief*: They might deny that people are really that happy, perhaps questioning whether people with kidney failure are giving honest answers when asked how happy they are, or believing that the DA developers have an agenda to promote.

(2) *Believe others but not self*: They might accept that most people with kidney failure are happy but not believe that *they* would be so happy.

(3) *Total acceptance*: They might believe that they, like the average person, would largely adapt to having kidney failure.

If they respond in the first manner, with complete disbelief, then we can confidently conclude that we have *not* adequately de-biased them from affective forecasting errors. In this case, we ought to do more to convince them or find yet other ways to debias them.

But what about the second type of response above, where a decision maker holds an accurate belief about how an average person responds to a given circumstance, but is convinced that he will respond differently? For any given individual, it is impossible to know whether the decision maker is right or wrong. Some people really

are made miserable by circumstances to which the average person can adapt. The same goes for the third response above. A person who believes that they will adapt to a given circumstance because most people adapt to the circumstance may *still* be making an affective forecasting error. For all we know, they won't adapt as much as the average person. In either of these cases, it is impossible to tell whether a specific individual has correctly predicted their response to a given circumstance.

Therefore, to judge de-biasing efforts, DA developers need to assess *aggregate responses*. If most decision makers believe they will be the exception to the rule, then the DA developer has not de-biased them. Ideally, the forecasts of decision makers will map, as a group, onto the actual reports of people experiencing the circumstances in question. Better yet, longitudinal studies could verify which ways of informing people about circumstances are best at mapping onto how they will actually respond.

In short, the evaluation of DAs should expand to test, in longitudinal studies, whether people who receive the DA are able to predict the emotional and non-emotional consequences of decision-relevant outcomes.

3. The Happiness Criterion

The flaws identified with these first two alternative criteria raise the possibility that our problem – of identifying when a third party has improved someone's decision– can be solved by resorting to a happiness criterion. Specifically, we could test whether people who receive a DA are happier than those who do not. This criterion is justified on the grounds that people don't know what makes themselves happy or unhappy, so a third party should figure out what does make them happy, and find a way to convince them to make decisions that maximize their happiness (P Ubel, 2009). This criterion differs from the second criterion – the forecasting criterion – by judging decisions after

the fact, rather than before. We know a decision is better based on the consequences of the decision, rather than the process that led to the decision.

The happiness criterion has many important strengths. All else equal, people generally want to be happy rather than unhappy, preferring positive moods to negative ones. Yet people do not always manage, even when well-informed and un-coerced, to make decisions that maximize these aspects of well-being (P. A. Ubel, 2006). Thus, it would seem to be a good thing for DA's to protect people from making decisions that reduce their happiness.

But the happiness criterion suffers from two major weaknesses. First, experts do not agree on how to define happiness. Some define happiness narrowly, as the balance of positive and negative affect (Bentham, 1907; Kahneman, Wakker, & Sarin, 1997). By this definition, happiness is quantifiable, and can be used to judge the impact of specific circumstances, or even decisions, on people's emotional well-being.

But this hedonic view of happiness strikes many people as being too narrow (Griffin, 1989; G Loewenstein & Ubel, 2008). People care about many aspects of their lives, beyond their moment-to-moment mood. For instance, they care about freedom, for freedom's sake, preferring to trade off some amount of happiness to increase their freedom. In addition, people care about opportunities and capabilities independent of how any limits on opportunities or capabilities influence their mood (Sen, 2004). Thus, for example, even when people recognize that their happiness will not be significantly reduced by a loss of income or by a new disability, most will nonetheless desire to maintain their income and their physical functioning (Damschroder, Zikmund-Fisher, & Ubel, 2005). Amartya Sen, in fact, contends that capabilities matter in large part because people are so good at adapting, emotionally, to unjust circumstances (Sen, 2004). Slavery would not be tolerable even if slaves were happy. Kidney failure would not be inconsequential just because people with kidney failure managed to adapt.

In short, and DAs should be evaluated to see if they increase people's overall sense of well-being, but we should also be aware of how DAs influence freedom and capabilities. We cannot assume that if a DA improves people's moods, it has therefore improved their decisions. Nor can we assume that if a DA reduces people's happiness, it has therefore influenced them to make bad decisions. Sometimes decision-makers make decisions solely to promote *other people's* interests, occasionally sacrificing their happiness for the sake of others. We would not want to call these decisions misguided.

4. The Invariance Criterion

DAs could also be evaluated by the standard of invariance – the idea that decisions should not change when the pros and cons of the decision alternatives remain the same. By this criterion, if I favor surgery A when I learn it has a 90% survival rate, then I should also favor it when I discover it has a 10% mortality rate—because a 90% survival rate is equivalent to a 10% mortality rate. Similarly, if I decide to take a medication with a 3-in-100 chance of migraines, I should not change my mind when I discover that the risk is 30-in-1000.

My colleagues and I have had success developing several methods for eliminating the influence of these decisional inconsistencies in health care DAs. For example, we discovered that graphical representations of probability information can reduce the inappropriate influence of anecdotes, what I refer to in my medical practice as "the Aunt Millie problem." Like most clinicians, I have encountered patients who reject plausible treatment alternatives out of hand because of something they heard from a friend or relative. Such encounters suggest that the way people feel about risks and benefits can be influenced by anecdotal information.

We explored this phenomenon in a survey of prospective jurors in Philadelphia (P. A. Ubel, Jepson, & Baron, 2001). We asked people to imagine that they had chest

pains from coronary artery disease, and that there were two treatment alternatives to choose from: bypass surgery, which had a 75% chance of curing their chest pains but which required open-heart surgery and prolonged recovery period; or balloon angioplasty, which had only a 50% chance of curing their chest pains but was a much less arduous procedure. We illustrated this choice with a series of uninformative anecdotes, relaying the stories of hypothetical people who had received each treatment and had either experienced a cure of their chest pain or had not experienced a cure.

Our study involved an experimental manipulation of the number and balance of anecdotes for each treatment alternative. One group of participants received *balanced anecdotes*, with two testimonials about each treatment – one from a person who got better and one from someone who did not. Another group received *statistically reinforcing anecdotes*: four testimonials from bypass patients, three of whom had gotten better and one who had not (thus, mirroring the 75% success rate of the treatment).

It is important to keep in mind that the anecdotes were uninformative. They did not illustrate anything about the treatment alternatives that people had not already been told. They simply relayed stories of treatment success or failures, and we had already informed them of the success rates of each treatment. Nevertheless, people's hypothetical treatment choices were significantly influenced by the anecdotes they encountered, with 30% of people receiving balanced anecdotes choosing bypass surgery versus 44% of those receiving statistically reinforcing anecdotes. Receiving a larger number of positive anecdotes about bypass surgery increased people's willingness to choose this treatment, even though those anecdotes told people nothing about the treatments that they didn't already know.

In a follow up study, we discovered that we could reduce the influence of anecdotes by providing graphical representations of the success rates of the two treatments (Fagerlin, Wang, & Ubel, 2005). We randomized participants so that half of

them received a pictorial representation of the success rates alongside the prose description. (The pictographs that we used are illustrated in Figure 1.) We found that the influence of anecdotes was eliminated by this pictograph. Regardless of whether we included "balanced anecdotes" or "statistically reinforcing anecdotes," approximately 40% of people chose bypass surgery. In other words, the influence of anecdotes was eliminated when the statistical information was supported by pictographs.

DAs should be judged for invariance. Two DAs that lead to equal comprehension of a specific decision can also lead to different decisions, if they introduce any of a number of decisional biases. Therefore, in judging DAs we should test for such biases, and when they are present we should develop ways to eliminate them.

5. Correlational Validity

All else equal, a woman with a high risk of breast cancer should be more interested in taking Tamoxifen than a woman with a moderate risk, who, on average, should be more interested than someone with a low risk. This is a standard for judging DAs that I refer to as correlational validity. If varying the risk-benefit ratio of a choice has no influence on people's decisions to choose that alternative, then DA developers have to worry that their DA is failing to make the tradeoffs clear, or is biasing people's choices.

Surprisingly, this standard is not generally used to evaluate healthcare DAs. In part, I expect this oversight has occurred because the medical decision making community has been so firmly wedded to the knowledge model of decision making, that they have not felt much need to test whether people are applying their knowledge rationally. In addition, I expect this criterion has been ignored because it doesn't provide clear guidance about what constitutes a good or bad decision, since there is no way to judge what the appropriate correlation should be between the risk/benefit ratio of an alternative and people's decisions. For example, suppose there's a correlation of 0.1 in

breast cancer risk and interest in Tamoxifen among women exposed to a given DA. Suppose there is a 0.3 correlation among women exposed to an alternative DA. Is either of these the correct correlation? We might feel confident that a DA that leads to no correlation is flawed, but can we be convinced that one of these DAs is better than the other?

In summary, DAs should be tested for correlational validity, and if the correlation is unacceptably low (which is a judgment call), then the DA should be revised to better highlight the risk/benefit trade-offs.

6. Time to Process the Decision

In medical practice, it is relatively common for men to receive diagnoses of prostate cancer at the same clinic visit in which a urologist helps them decide how to treat their prostate cancer. More often than not, men leave such visits deciding for or against surgical intervention. Sometimes they choose radiation treatment, but often they haven't even had the chance to meet with a radiation oncologist.

There is increasing evidence in the decision science literature that time is a crucial element of optimal decision making. Although controversial, some studies suggest that unconscious deliberation can improve people's decision making (Dijksterhuis, Bos, Nordgren, & van Baaren, 2006). Such deliberation takes time. There's also ample evidence that people make different decisions when in hot emotional states versus cold (G. Loewenstein, 1999). With so much information to process, and so many options to consider, it hardly seems plausible that a person who has just found out he has cancer would be able to make a good decision quickly.

Thus, DA evaluation should be broadened to include an assessment of whether people had enough time to process the decision.

7. Adherence

Some decisions are “one and done” affairs—choose surgery over, say, chemo and you will receive the surgery and your decision will be irreversible. But many decisions are not so final. A patient who decides to take a cholesterol pill, for example, faces that decision every day. A person who decides to save more money, and reduce entertainment expenditures, still faces the temptation to splurge on a nice vacation.

A good DA, then, will not only help people make a decision—about whether to take a pill or a vacation—but will also help them *stick* with the decision. Such DAs should therefore be evaluated for how frequently people adhere to the decisions they make.

Conclusion

I have laid out a few criteria by which to determine whether a structured DA has helped people make decisions that reflect any underlying preferences they have. None of the criteria are on their own sufficient to prove that a DA has led to unbiased decisions. Thus, DA developers need to use careful judgment in applying these criteria to any existing DA, recognizing tradeoffs between the potential attributes of a DA and the various outcomes that decision makers care about. When viewed as a whole, these criteria should give decision counselors a much better idea of when they are helping decision makers. These expanded criteria certainly provide a better idea of the strengths and weaknesses of DAs than do the knowledge and satisfaction-based criteria that have dominated the field to date.

References

Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior & Human Decision Processes, 70*(1), 1-16.

Bekker, H., Thornton, J. G., Airey, C. M., & al., e. (1999). Informed decision making: An annotated bibliography and systematic review. *Health Technology Assessment, 3*, 1-156.

Bentham, J. (1907). *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.

Damschroder, L. J., Zikmund-Fisher, B. J., & Ubel, P. A. (2005). The impact of considering adaptation in health state valuation. *Social Science & Medicine, 61*(2), 267-277.

Day, R. (2001). Quality of life and tamoxifen in a breast cancer prevention trial: A summary of findings from the NSABP P-1 study. *Annals of the New York Academy of Sciences, 949*, 143-150.

Diener, E., & Seligman, M. E. P. (2004). Beyond money: Toward an economy of well-being. *Psychological Science in the Public Interest, 5*(1), 1-31.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: the deliberation-without-attention effect. *Science, 311*(5763), 1005-1007.

Fagerlin, A., Wang, C., & Ubel, P. (2005). Reducing the influence of anecdotal reasoning on people's health care decisions: Is a picture worth a thousand statistics? *Medical Decision Making, 25*(4), 398-405.

Gilbert, D., Killingsworth, M., Eyre, R., & Wilson, T. (2009). The surprising power of neighborly advice. *Science, 323*(5921), 1617-1619.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. (Eds.). (1996). *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press.

Green, P., & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5*, 103-123.

Griffin, J. (1989). *Well Being. Its meaning, measurement, and moral importance.*: Clarendon.

Holmes-Rovner, M., Kroll, J., Schmitt, N., Rovner, D. R., Breer, M. L., Rothert, M. L., et al. (1996). Patient satisfaction with health care decisions: The satisfaction with decision scale. *Medical Decision Making, 16*(1), 58-64.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics, 112*(2), 375-405.

Kennedy, D. M. (2003). On what basis should the effectiveness of decision aids be judged? *Health Expectations, 6*, 255-268.

Loewenstein, G. (Ed.). (1999). *A visceral account of addiction*. Cambridge, UK: Cambridge University Press.

Loewenstein, G., & Ubel, P. (2008). Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics, 92*(8-9), 1795-1810.

McNeil, B. J., Pauker, S. G., Sox, H. C., Jr., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine, 306*(21), 1259-1262.

Molenaar, S., Sprangers, M. A., Postma-Schuit, F. C., Rutgers, E. J., Noorlander, J., Hendriks, J., et al. (2000). Feasibility and effects of decision aids. *Medical Decision Making, 20*(1), 112-127.

Nelson, W. L., Han, P. K., Fagerlin, A., Stefanek, M., & Ubel, P. A. (2007). Rethinking the objective of decision aids: A call for conceptual clarity. *Medical Decision Making, 27*(5), 609-618.

O'Conner, A. M. (1993). Decisional conflict. . In G. K. McFarland & E. A. McFarlane (Eds.), *Nursing diagnosis and intervention.* (pp. 468-478). St Louis.: Mosby.

O'Conner, A. M., Fiset, V., Rostom, A., Tetroe, J., Entwistle, V., Llewellyn-Thomas, H. A., et al. (2004). Decision aids for people facing health treatment or screening decisions (Cochrane Review). In *The Cochrane Library* (Vol. 4). Chichester, UK: John Wiley & Sons, Ltd.

O'Conner, A. M., Jacobsen, M. J., & Stacey, D. (2002). An evidence-based approach to managing women's decisional conflict. *Journal of Obstetric, Gynecologic and Neonatal Nursing, 31*, 570-581.

O'Connor, A. M. (1995). Validation of a decisional conflict scale. *Medical Decision Making, 15*(1), 25-30.

O'Connor, A. M., Entwistle, S. D., Llewellyn-Thomas, H., Rovner, D., Holmes-Rovner, M., Tait, V., et al. (2003). Decision aids for people facing health treatment or screening decisions. Issue 2. 2003, from http://www.cochrane.org/cochrance/revabstr/ab001431.htm

O'Connor, A. M., Fiset, V., DeGrasse, C., Graham, I. D., Evans, W., Stacey, D., et al. (1999). Decision aids for patients considering options affecting cancer outcomes: Evidence of efficacy and policy implications. *Journal of the National Cancer Institute. Monographs*(25), 67-80.

Riis, J., Loewenstein, G., Baron, J., Jepson, C., Fagerlin, A., & Ubel, P. A. (2005). Ignorance of hedonic adaptation to hemodialysis: a study using ecological momentary assessment. *J Exp Psychol Gen, 134*(1), 3-9.

Sen, A. (2004). Capabilities, lists, and public reason: Continuing the conversation. *Feminist Economics, 10*(3), 77-80.

Smith, D., Loewenstein, G., Jankovich, S., Jepson, C., Feldman, H., & Ubel, P. (2008). Mispredicting and misremembering: patients with renal failure overestimate improvements in quality of life after a kidney transplant. *Health Psychology, 27*(5), 653-658.

Smith, D. M., Sherriff, R. L., Damschroder, L., Loewenstein, G., & Ubel, P. A. (2006). Misremembering colostomies? Former patients give lower utility ratings than do current patients. *Health Psychology, 25*(6), 688-695.

Torrance, G. (1976). Social preferences for health states: An empirical evaluation of three measurement techniques. *Socioeconomic Planning Science, 10,* 129-136.

Torrance, G. W., Thomas, W. H., & Sackett, D. L. (1972). A utility maximization model for evaluation of health care programs. *Health Services Research., 7*(2), 118-133.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science., 211*(4481), 453-458.

Ubel, P. (2009). *Free Market Madness: Why Human Nature is at Odds with Economics--and Why it Matters*. Boston: Harvard Business Press.

Ubel, P., Loewenstein, G., & Jepson, C. (2005). Disability and sunshine: Can predictions be improved by drawing attention to focusing illusions or emotional adaptation? *Journal Experimental Psychology: Applied, 11*(2), 111-123.

Ubel, P. A. (2006). *You're stronger than you think: Tapping into the secrets of emotionally resilient people.* New York: McGraw-Hill.

Ubel, P. A., Jepson, C., & Baron, J. (2001). The inclusion of patient testimonials in decision aids: Effects on treatment choices. *Medical Decision Making, 21*(1), 60-68.

Ubel, P. A., & Loewenstein, G. (1997). The role of decision analysis in informed consent: choosing between intuition and systematicity. *Social Science & Medicine, 44*(5), 647-656.

Ubel, P. A., Loewenstein, G., Schwarz, N., & Smith, D. (2005). Misimagining the unimaginable: The disability paradox and healthcare decision making. *Health Psychology, 24*(4 Supplement), S57-S62.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior (2nd rev. ed.).* Princeton, NJ: Princeton University Press.

Wilson, T., Meyers, J., & Gilbert, D. (2003). How happy was I anyway? a retrospective impact bias. *Social Cognition, 21*, 407-432.

Woloshin, S., Schwartz, L. M., Moncur, M., Gabriel, S., & Tosteson, A. N. (2001). Assessing values for health: Numeracy matters. *Medical Decision Making, 21*(5), 382-390.

Zikmund-Fisher, B., Fagerlin, A., Roberts, T., Derry, H., & Ubel, P. (2008). Alternate methods of framing information about medication side effects: Incremental risk versus total risk occurence. *Journal of Health Communication, 13*(2), 107-124.

Table 1: Criteria available for evaluating whether decision aid has improved decisions

Standard criteria used in health care

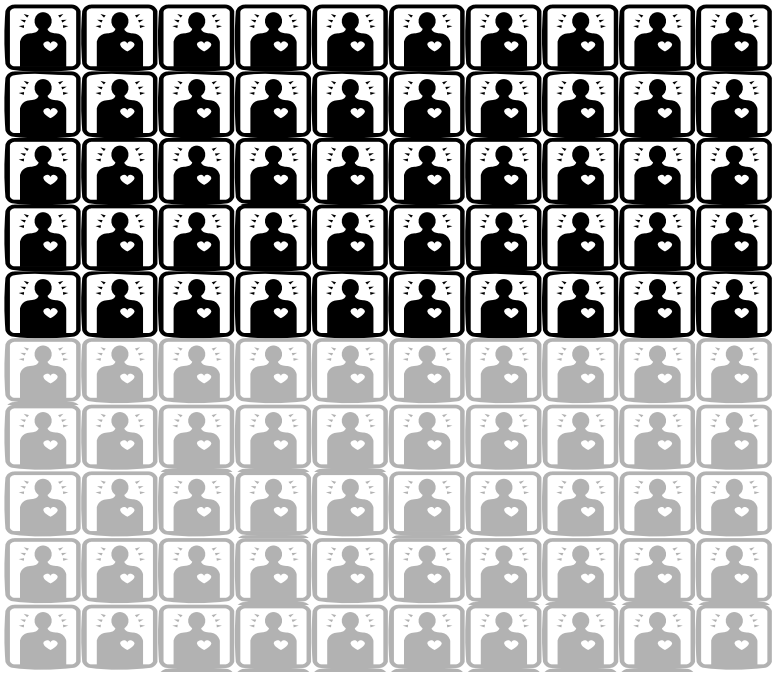
1. Increase in knowledge
2. Satisfaction with decision
3. Reduction in decisional conflict

New criteria discussed in this chapter

1. Maximization of expected utility
2. Reduction of mispredictions: e.g. accurate beliefs about consequences
3. Increased happiness: i.e. moment to moment moods
4. Invariance: decision less susceptible to non-normative influences
5. Correlational validity: decision shifts appropriately with change in risks/benefits
6. Time: decision maker has adequate time to process alternatives
7. Adherence: decision maker follows through on decision

Figure 1: Pictographs used to communicate cure rate of bypass surgery and balloon angioplasty.

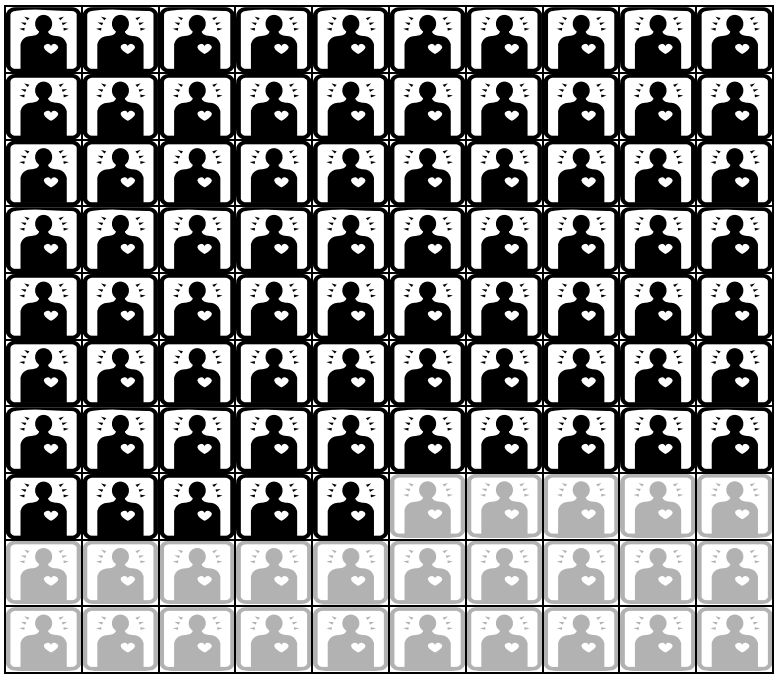
Success Rate of Balloon Angioplasty



Successfully cured of angina

Not successfully cured of angina

Success Rate of Bypass Surgery



Successfully cured of angina

Not successfully cured of angina