

Are patient decision aids effective? Insight from revisiting the debate between correspondence and coherence theories of judgment

Victoria A. Shaffer* and Lukas Hulsey

Department of Psychology

Wichita State University

Abstract

Research endeavors to determine the effectiveness of patient decision aids (PtDAs) have yielded mixed results. The conflicting evaluations are largely due to the different metrics used to assess the validity of judgments made using PtDAs. The different approaches can be characterized by Hammond's (1996) two frameworks for evaluating judgments: correspondence and coherence. This paper reviews the literature on the effectiveness of PtDAs and recasts this argument as a renewed debate between these two meta-theories of judgment. Evaluation by correspondence criteria involves measuring the impact of patient decision aids on metrics for which there are objective, external, and empirically justifiable values. Evaluation on coherence criteria involves assessing the degree to which decisions follow the logical implications of internal, possibly subjective, value systems/preferences. Coherence can exist absent of correspondence and vice versa. Therefore, many of the seemingly conflicting results regarding the effectiveness of PtDAs can be reconciled by considering that the two meta-theories contribute unique perspectives. We argue that one approach cannot substitute for the other, and researchers should not deny the value of either approach. Furthermore, we suggest that future research evaluating PtDAs include both correspondence and coherence criteria.

Keywords: correspondence, coherence, patient decision aids, medical decision making.

1 Introduction

In 1996, Kenneth Hammond argued that the field of judgment and decision making has taken two distinct paths, describing those two paths as meta-theories used in the evaluation of judgments. In the correspondence meta-theory, judgments are compared to objective facts with the explicit goal of evaluating the empirical accuracy of the judgments. In contrast, the coherence meta-theory assesses the logical consistency of judgments. In this framework, judgments are evaluated by the degree to which they are similar to solutions given by logical, mathematical, or statistical theories. Judgments are not compared to an external "true" value; instead judgments are compared to a normative rule, often mathematical in nature. While correspondence focuses on the accuracy of judgments, coherence focuses on the rationality of judgments.

*Address: Victoria A. Shaffer, Wichita State University, 1845 Fairmount St., Wichita, KS 67260-0034. Email: victoria.shaffer@wichita.edu.

2 Application to patient decision aids

Hammond's meta-theories have been applied to social policy in many domains including medicine. There are many applications within the medical field; however, this article pertains specifically to treatment choices made by patients. For many treatment decisions, there is a dominant treatment choice, or a single "best" therapeutic action. This choice is easy because the treatment has a clear advantage — effectiveness is high and/or side effects are low. These treatment choices do not necessitate much assistance. However, there exist a number of medical conditions for which no single treatment choice dominates. These are the cases in which the patient must make trade-offs between different dimensions or features of treatment options. For example, risk and potential benefit are often positively correlated, and a choice must depend upon how an individual values the potential benefits and harms. Currently there is no best therapeutic course of action for menopause, and women are commonly asked to choose whether or not to begin hormone replacement therapy

(HRT). In addition to alleviating some of the symptoms of menopause (weight gain, hot flashes, poor memory, etc.), HRT also helps to protect the body from osteoporosis. However, HRT is also associated with an increased risk of breast cancer, heart disease and stroke. Therefore, each individual patient must weigh the increased health risks against the benefits, protection from osteoporosis and reduction in symptoms. Other similar examples include treatment for menorrhagia, ischaemic heart disease, atrial fibrillation, abnormal uterine bleeding, prostate cancer, and breast cancer.

These types of treatment choices are very difficult for patients and their physicians cannot guide them to the “right” answer. Therefore, a class of decision aids has been developed to help patients make difficult treatment choices between two or more options. Patient decision aids (PtDAs) are designed to educate patients on the various treatment choices and their associated risks and benefits, help patients understand the values they place on those risks and benefits, and improve communication between patients and doctors. The emphasis on improving patient decision making evolved during the shift in patient care philosophy from informed consent to informed choice or shared decision making. PtDAs were heralded as a way to improve informed choice, reduce geographic differences in treatment choices, combat the rise of consumerism in medicine, and increase patient satisfaction with the treatment process (O’Conner et al., 2003).

The number of PtDAs steadily increases each year, and the medical community has recently begun to evaluate the effectiveness of these programs. In 2003, the Cochrane Collaboration Patient Decision Aid Review Group published an extensive evaluation of the effectiveness of PtDAs. In 2006, the Society for Medical Decision Making held the International Patient Decision Aids Standards (IPDAS) symposium, which was designed to debate whether PtDAs were the best way to improve patient decision making. Subsequently, the journal *Medical Decision Making* dedicated an entire issue to the future of shared decision making in 2007, which contained a focus on evaluating the effectiveness of PtDAs (Helfand, Barnato, & Holmes-Rovner, 2007). What has emerged is a set of conflicting opinions about the value of PtDAs and, more basically, about the criteria upon which PtDAs should be evaluated. Reviewing this debate shows that its arguments parallel those between the correspondence and coherence theorists of decision making. Hence, the goal of this paper is to re-cast this discussion in light of these two meta-theories in the hope of learning lessons from a previous incarnation of this debate.

3 The debate about the effectiveness of patient decision aids

The Cochrane Collaboration Patient Decision Review Group (2003) summarized the results of 35 randomized controlled trials that evaluated the effectiveness of PtDAs in comparison with control groups receiving the usual care or simpler decision aids. Following the Cochrane Review, the IPDAS symposium was created to debate whether or not patient decision aids are the best way to improve clinical decision making. A vote after the symposium revealed that medical professionals are divided in their opinion of the effectiveness of PtDAs with about half believing PtDAs to be effective in improving clinical decision making thus favoring their widespread implementation and half believing PtDAs do not improve clinical decision making thus opposing their implementation (Holmes-Rovner et al., 2007). Although a number of other issues were addressed, a major portion of the debate focused on the criteria for evaluating the success of PtDAs. Wendy Nelson, of the HHS National Cancer Institute, described the current mixed bag approach to evaluating PtDAs: “What is a quality decision, and what constitutes quality decision making? Because we lack an agreed-upon definition of quality, proponents of decision aids have operationalized quality in a wide variety of ways” (Holmes-Rovner et al., 2007, p. 602).

Researchers evaluating PtDAs determine the effectiveness of a given decision aid by assessing whether or not it improves patient decision making. There have been two distinct approaches to operationalizing improved decision making, one that roughly captures the coherence approach and one that more closely approximates the correspondence approach. Researchers adopting the coherence approach often ask whether the use of PtDAs improves the congruence between patients’ stated values and their treatment choices (O’Connor et al., 2007; O’Connor et al., 2003). In contrast, researchers adopting the correspondence approach typically ask whether the use of PtDAs results in increased quality of life (McCaffery, Irwig, & Bossuyt, 2007). These two methods frequently lead to different conclusions about the effectiveness of PtDAs. We argue that the two sides in the PtDA controversy represent a renewed debate about correspondence versus coherence criteria. The following sections will discuss how the correspondence and coherence theories have been applied to the area of PtDAs and the benefits and difficulties associated with adopting each approach.

4 The correspondence approach to evaluating patient decision aids

Hammond's 1955 article, "Probabilistic functioning and the clinical method," marked the beginning of the correspondence approach to decision making. Although formally articulated by Hammond in the 1950s, the correspondence metatheory was inspired by Egon Brunswik's general theory of cognition under uncertainty. The correspondence approach focuses on aligning human judgment with empirical accuracy. Typically, correspondence theorists have evaluated decision making by comparing human judgment to an external gold standard — often objective facts (Hammond, 1996).

In the context of patient decision making, the logical application of the correspondence approach would be to examine whether patient treatment choices corresponded to the best treatment choices. However, PtDAs have been developed for treatment choices for which there exists no single best therapeutic action. Therefore, this obvious correspondence criterion is eliminated. Instead, "correspondence" theorists evaluating PtDAs argue for the implementation of other gold standards such as survival, function and well-being. More "accurate" choices lead to a greater chance of survival, greater post-treatment function and improved quality of life. Health outcomes have been evaluated using general health measures such as the 12-item Short-form Health Survey or measures that assess condition-specific health outcomes (O'Connor et al., 2003). However, very few studies have used these measures. Out of the 35 randomized controlled trials of PtDAs reported in the Cochrane Review, only seven studies measured general or condition specific health outcomes. Five of these seven studies found no significant differences in health outcomes for patients who used decision aids. Furthermore, these metrics appear to have been subsequently abandoned. The IPDAS symposium, designed to evaluate the effectiveness of PtDAs, defined a good patient decision as "one that is informed and consistent with the patient's values" (Holmes-Rovner et al., 2007, p. 603). This definition is generally inconsistent with the correspondence approach, and, therefore, the majority of recent research on PtDAs has not included correspondence criteria. In fact, a recent meta-analysis of the effectiveness of PtDAs did not include any these outcome measures, abandoning the correspondence approach in favor of coherence (O'Connor et al., 2007).

There are difficulties associated with the correspondence approach that may have led some researchers to favor a coherence approach to evaluating PtDAs. First, the correspondence measures are difficult to evaluate in the short term. Metrics such as survival and long-term health outcomes require many years of follow-up data to be collected after the initial randomized controlled trial. There-

fore, if these were the only measures adopted, it may be difficult to conduct timely research. In addition, some researchers have adopted the position that decision-making should be evaluated *before* the outcome of the decision is known. In support of this position, Baron and Hershey (1988) documented the outcome bias in evaluating decisions; participants rated the quality of the decision and the competence of the decision maker more favorably after a positive outcome than a negative outcome. Furthermore, Caplan and colleagues (1991) found a negative relationship between outcome severity and judgments of the appropriateness of care by anesthesiologists.

5 The coherence approach to evaluating patient decision aids

Due to the lack of a clear gold standard and the difficulties associated with implementing the correspondence approach, a number of researchers have opted to evaluate the efficacy of treatment decisions by using coherence criteria in lieu of correspondence criteria. This approach to the study of decision making developed in parallel to Hammond's correspondence approach. Ward Edwards is largely credited with beginning this movement in 1954 with his article on "The theory of decision making." Edwards' research examined the rationality of human judgment under uncertainty using probability theory as the cornerstone of rationality. Edwards evaluated judgmental ability by comparing human judgment to mathematical models — as opposed to empirical accuracy. Edwards focused on a specific type of probability calculus — Bayes' Theorem. Later coherence theorists also adopted Expected Utility Theory (EUT) and its variants and their axioms as additional yardsticks for rationality.

Most research on PtDAs comes from the coherence approach. Because there is no objective gold standard or single best treatment, most research examining the effectiveness of PtDA asks whether the aid helps the patient to make a coherent choice (O'Connor et al., 2007; O'Connor et al., 2003). In fact, two explicit goals of many PtDAs are to help patients clarify their values for various health states and make treatment choices that are congruent with those values. For example, a patient who values treatment effectiveness should be more likely to choose a treatment with a greater success rate. However, if she values quick recovery times or pain minimization, then she should be more likely to choose treatments that match this. Many PtDAs include values clarification exercises to help patients articulate their preferences; 30 out of 51 PtDAs included in a recent meta-analysis incorporated some explicit values clarification exercise (O'Connor et al., 2007). These exercises can include rating different possible health outcomes, forced choice

tasks where patients choose between options with a single difference, and time tradeoff tasks where the patient is asked to decide how many years of perfect health they would trade in exchange for a certain number of years in some less desirable health state. The goal of these exercises is to determine how an individual patient values the relative risks and benefits associated with the different potential outcomes. Therefore, the coherence approach to determining the efficacy of a PtDA is to examine whether treatment choices made by patients using PtDAs are congruent with their stated values (Holmes-Rovner et al., 1999; Lerman et al., 1997; O'Connor et al., 1999)

Like the correspondence framework, researchers have argued that the coherence approach to evaluating patient judgments has some weaknesses. Specifically, evaluating the rationality of a treatment decision by its congruence to elicited preferences for health states relies on the validity of the system used to elicit those preferences. There are some reasons to question the validity of values clarification exercises commonly built into PtDAs (Nelson, Han, Fagerlin, Stefanek, & Ubel, 2007). These exercises assume that preferences for health states are stable and that patients are accurately able to predict preferences for future health states. However, research suggests that these assumptions could be questioned.

The coherence criterion of value congruence relies on the assumption that patients have measurable and stable preferences for health states and treatment choices. This assumption has been hotly contested by decision making researchers since the 1970s. A large body of research has demonstrated that preferences appear to be constructed at the time of judgment and are sensitive to the mode of elicitation and to framing effects (Slovic, 1995; Slovic & Lichtenstein, 1983; Tversky & Kahneman, 1981). Therefore, in the context of a PtDA, preferences for health states may be subject to framing effects, and values clarification exercises may help to determine preferences — not discover them. This again calls into question the use of value congruence as a yardstick for rationality.

Additionally, the validity of the values that patients assign to health states is also questionable. Even if we assume that patient preferences for health states are measurable and stable, values clarification exercises rely on the assumption that patients are accurately able to predict their utility for future health states. For example, imagine a man diagnosed with localized prostate cancer needs to choose between two approaches to dealing with the disease: watchful waiting or prostatectomy. In a very small proportion of men, a prostatectomy may be effective in preventing a metastasis of the cancer. However, the side effects of prostatectomy, impotence and incontinence, occur in a large proportion of men. In contrast, there are no side effects, other than perhaps in-

creased stress, to watchful waiting, and the cancer will not metastasize in the majority of cases — although this number is smaller than with a prostatectomy. Therefore, the patient's choice would be based on his utilities for the potential health states. Which would be more difficult to live with: a small increase in the likelihood that the cancer would metastasize or the risk of impotence/incontinence? In order to make the decision, the patient has to be able to successfully predict how he would feel in these previously unexperienced health states. Much research has shown that people are unable to predict their future feelings (Damschroder, Zikmund-Fisher, & Ubel, 2005; Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Riis et al., 2005; Ubel, Loewenstein, & Jepson, 2005a; Ubel, Loewenstein, Schwarz, & Smith, 2005b; Wilson & Gilbert, 2005). Wilson and Gilbert (2005) argued that people are unsuccessful at "affective forecasting" because they mispredict their emotional reactions to future events. In particular, people fail to anticipate how readily they will adapt to emotionally difficult situations. This phenomenon has been demonstrated with affective predictions for various health states as well (e.g., Brickman, Coates, & Janoff-Bulman, 1978; Ubel et al., 2005b). For example, Sackett and Torrance (1978) found that patients with chronic health conditions, such as dialysis, have a greater quality of life than is predicted by people without a chronic illness. Therefore, the basic premise upon which values clarification exercises are designed — namely that patients can predict their utilities for future health states — is called into question.

The variables we have summarized in the preceding sections do not represent an exhaustive list of the measures used to evaluate PtDAs. There are additional variables that cannot be easily classified as fitting into either approach or that represent a hybrid of the two approaches. For example, researchers have examined a number of psychological measures such as anxiety, depression, regret, and decisional conflict. Some take a reduction in these psychological variables as evidence for the effectiveness of PtDAs (O'Connor et al., 2003; O'Connor et al., 2007). Similarly, some researchers argue that other non-psychological measures such as persistence with a chosen treatment choice, increased participation in decision making, and reduced indecision after PtDA provide evidence that PtDAs are effective tools. Finally, some researchers examine whether the use of PtDAs improves the accuracy of risk perception, which is often determined by normative solutions such as Bayes' Theorem.

Although some measures do not fit neatly into the correspondence/coherence distinction, we find it a useful classification tool for PtDA evaluation metrics. Correspondence measures speak to the empirical accuracy of judgments, while coherence measures speak to the logical consistency of judgments. Measures from these two

categories directly evaluate the quality of patient judgments. Variables outside this classification system would at best be considered indicators of good judgment.

6 Conclusions: Suggestions for future PtDA evaluation

PtDAs have been evaluated from both a coherence and correspondence perspective, and the two approaches have yielded conflicting conclusions. Research from the coherence approach has concluded that PtDAs help patients to make more coherent decisions, i.e., the decisions reached after consulting these aids are generally more consistent with patients' values (O'Connor et al., 2007; O'Connor et al., 2003). In contrast, research from the correspondence approach has been less positive. Although PtDAs have been shown to improve risk perception, there is little evidence that PtDAs increase general or condition-specific health outcomes, quality of life, or survival (O'Connor et al., 2003). Although the two approaches lead to different conclusions (coherence theorists would argue that PtDAs are effective, correspondence theorists would argue they are not), Hammond would contend that the two approaches do not necessarily need to compete (Hammond, 1996). Instead, he suggests that the two perspectives can and should peacefully co-exist because they complement each other. The fact that the two meta-theories have different approaches and do not measure the same constructs implies that they both contribute unique information. Therefore, one approach cannot substitute for the other, and researchers should not deny the value of either approach. "Complementarity has its own virtues, however, for each theory, together with its associated methodology, affords the researcher an opportunity to explore the nature of human judgment in ways that supplement one another. And when both are applied, they help us overcome the limitations of each alone and thus afford us knowledge we would not obtain otherwise" (Hammond, 1996, p. 229). Therefore, we argue that researchers should be cognizant of the approach their research adopts and should recognize that the two approaches have different, but not contradictory, goals. To this end, PtDA should continue to be evaluated from both the correspondence and coherence approaches. It is equally important to learn about contributions that PtDAs make to both internal consistency and long term health outcomes. Thus we are advocating for continued research from both approaches.

However there is currently an unequal contribution between the two approaches, with the majority of the research evaluating the effectiveness of PtDAs from the coherence approach. Unfortunately, we do not foresee this discrepancy between the contributions of the two ap-

proaches improving in the next several years. In fact, the field appears to be increasingly dominated by the coherence approach. The recent IPDAS symposium on this issue advocated for evaluating the effectiveness of PtDAs from the coherence approach, with little regard for the correspondence approach (Holmes-Rovner et al., 2007). It was argued that PtDAs should be evaluated by whether or not they lead patients to a good decision, with the hallmark of a good decision being one that is consistent with a patient's values. We do not believe that coherence approaches should be abandoned; instead, we call for a commensurate increase in correspondence research. We recommend that all randomized controlled trials evaluating patient decision aids include both coherence and correspondence criteria.

References

- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54*, 569–579.
- Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology, 36*, 917–927.
- Caplan, R. A., Posner, K. L., & Cheney, F. W. (1991). Effect of outcome on physician judgments of appropriateness of care. *JAMA, 265*, 1957–1960.
- Damschroder, L. J., Zikmund-Fisher, B. J., & Ubel, P., A. (2005). The impact of considering adaptation in health state valuation. *Social Science & Medicine, 61*, 267–277.
- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 41*, 380–417.
- Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 75*, 617–638.
- Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review, 62*, 255–262.
- Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, and unavoidable injustice*. New York: Oxford University Press.
- Helfand, M., Barnato, A., & Holmes-Rovner, M. (Eds.). (2007). Shared decision making, decision aids, and risk communication [Special Issue]. *Medical Decision Making, 27*(5).
- Holmes-Rovner, M., Kroll, J., Rovner, D., Schmitt, N., Rothert, M., Padonu, G., & Talarczyk, G. (1999). Patient decision support intervention: Increased consistency with decision analytic models. *Medical Care,*

- 37, 270–284.
- Holmes-Rovner, M., Nelson, W. L., Pignone, M., Elwyn, G., Rovner, D. R., O'Connor, A. M., Coulter, A., & Correa-de-Araujo, R. (2007). Are patient decision aids the best way to improve clinical decision making? Report of the IPDAS symposium. *Medical Decision Making, 27*, 599–608.
- Lerman, C., Biasecker, B., Benkendorf, J. L., Kerner, J., Gomez-Caminero, A., Hughes, C., & Reed, M. M. (1997). Controlled trial of pretest education approaches to enhance informed decision-making for BRCA1 gene testing. *Journal of the National Cancer Institute, 89*, 148–157.
- McCaffery, K., Irwig, L., & Bossuyt, P. (2007). Patient decision aids to support clinical decision making: Evaluating the decision or the outcomes of the decision. *Medical Decision Making, 27*, 619–625.
- Nelson, W. L., Han, P. K. J., Fagerlin, A., Stefanek, M., & Ubel, P. A. (2007). Rethinking the objectives of decision aids: A call for conceptual clarity. *Medical Decision Making, 27*, 609–618.
- O'Connor, A. M., Bennett, C., Stacey, D., Barry, M., Col, N. F., Eden, K. B., Entwistle, V., Fiset, V., Holmes-Rovner, M., Khangura, S., Llewellyn-Thomas, H., & Rovner, D. R. (2007). Do patient decision aids meet effectiveness criteria of the international patient decision aids standards collaboration? A systematic review and meta-analysis. *Medical Decision Making, 27*, 554–574.
- O'Connor, A. M., Stacey, D., Entwistle, V., Llewellyn-Thomas, H., Rovner, D., Holmes-Rovner, M., Tait, V., Tetroe, J., Fiset, V., Barry, M., & Jones, J. (2003). Decision Aids for people facing health treatment or screening decisions *Cochrane Database of Systematic Reviews*. No.: CD001431.
- O'Connor, A. M., Wells, G. A., Tugwell, P., Laupacis, A., Elmslie, T., & Drake, E. (1999). The effects of an 'explicit' values clarification exercise in a women's decision aid regarding postmenopausal hormone therapy. *Health Expectations, 2*, 21–32.
- 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. *Journal of Experimental Psychology: General, 134*, 3–9.
- Sackett, D. L., & Torrance, G. W. (1978). The utility of different health states as perceived by the general public. *Journal of Chronic Diseases, 31*, 697–704.
- Slovic, P. (1995). The construction of preference. *American Psychologist, 50*, 364–371.
- Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review, 73*, 596–605.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458.
- Ubel, P. A., Loewenstein, G., & Jepson, C. (2005a). Disability and sunshine: Can predictions be improved by drawing attention to focusing illusions or emotional adaptation? *Journal of Experimental Psychology: Applied, 11*, 111–123.
- Ubel, P. A., Loewenstein, G., Schwarz, N., & Smith, D. (2005b). Misimagining the unimaginable: The disability paradox and health care decision making. *Health Psychology, 24*, 57–62.
- Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science, 14*, 131–134.