Published in final edited form as:

Med Care Res Rev. 2014 October ; 71(5 Suppl): 38S–64S. doi:10.1177/1077558713496321.

Complexity, Public Reporting, and Choice of Doctors: A Look Inside the Blackest Box of Consumer Behavior

Mark Schlesinger¹, David E. Kanouse², Steven C. Martino³, Dale Shaller⁴, and Lise Rybowski⁵

¹Yale University School of Public Health, New Haven, CT, USA

²RAND Health, Santa Monica, CA, USA

³RAND Health, Pittsburgh, PA, USA

⁴Shaller Consulting Group, Stillwater, MN, USA

⁵The Severyn Group, Ashburn, VA, USA

Abstract

Health care consumers often make choices that are imperfectly informed and inconsistent with their expressed preferences. Past research suggests that these shortcomings become more pronounced as choices become more complex, through either additional options or more performance metrics. But it is unclear why this is true: Consumer choice remains a “black box” that research has scarcely illuminated. In this article, we identify four pathways through which complexity may impair consumer choice. We examine these pathways using data from an experiment in which consumers (hypothetically) selected a primary care physician. Some of the loss of decision quality accompanying more complex choice sets can be explained by consumers’ skills and decision-making style, but even after accounting for these factors, complexity undermines the quality of decision making in ways that cannot be fully explained. We conclude by discussing implications for report designers, sponsors, and policy makers aspiring to promote consumer empowerment and health care quality.

Keywords

bounded rationality; choice; consumerism; experiment; decision styles; heuristics

We are built to make mistakes, coded for error.

—Lewis Thomas

Reprints and permissions: sagepub.com/journalsPermissions.nav

Corresponding Author: Mark Schlesinger, Yale University School of Public Health, Room 304 LEPH 60 College Street, New Haven, CT 06520, USA. mark.schlesinger@yale.edu.

Authors’ Note

This article was improved by comments from Andrew Parker, Rachel Grob, two anonymous reviewers, and the editors on this special issue.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Good decisions come from experience. Experience comes from making bad decisions.

—Mark Twain

Introduction

Health care consumers often make choices that are imperfectly informed and inconsistent with their expressed preferences (Barnes, Hanoch, Wood, Liu, & Rice, 2012; Faber, Bosch, Wollersheim, Leatherman, & Grol, 2009; Hibbard, Greene, Sofaer, Firminger, & Hirsh, 2012; Kling, Mullainathan, Shafir, Vermeulen, & Wrobel, 2012). Given the serious consequences of poor choices, policy makers in both public and private sectors have been motivated to try to enhance the decision-making process by providing more complete information, creating inducements to consider information more carefully, and offering support for consumers who desire assistance (Grob, Schlesinger, Davis, Cohen, & Lapps, 2013; McCormack, Schnaier, Lee, & Garfinkel, 1996; Sinaiko, Eastman, & Rosenthal, 2012).

Despite these investments—and the best intentions of policy makers—there is little evidence that efforts to enhance consumer decision making have been effective (French et al., 2009; Fung, Lim, Mattke, Damberg, & Shekelle, 2008; Sinaiko et al., 2012). This ineffectiveness is in large part because those seeking to improve consumer choices are essentially flying blind: Consumers' efforts to make sense of their health care choices remain a black box into which research has shined little light. We know from past studies that many consumers make suboptimal decisions even when information that could improve their choices is widely and freely available. This suggests that they devalue, ignore, misunderstand, or have difficulty using the information (Kling et al., 2012; Sinaiko & Hirth, 2011; Werner & Asch, 2005), but we know little about *why* this might be the case. And therefore we do not know how reports might be designed or disseminated to reduce these impediments to thoughtful use.

Recent studies have revealed one important clue. It appears that for both health plans and health care providers, the quality of decision making erodes as the choice set gets larger and more complex with respect to the amount and variety of information (Hanoch, Rice, Cummings, & Wood, 2009; Kanouse, Schlesinger, Shaller, Martino, & Rybowski, 2012). This pattern is consistent with research on decisions outside health care: Consumers have more difficulty choosing among a multitude of options than among fewer alternatives (Begley, 2011; Iyengar, Jiang, & Huberman, 2004). For policy makers to offer consumers more manageable health care choices and more efficacious assistance, it is essential to understand why larger choice sets result in less-than-ideal decisions.

To explore these matters, we identify four ways (causal pathways) in which more complex choice sets might undermine consumer decision making. We then draw on our own experimental research to explore the impact of each pathway on consumer choice of clinicians. Previously, we have established that consumers made worse choices selecting among physicians when presented with more complex performance metrics and larger choice sets (Kanouse et al., 2012). In this article, we assess the relative impact of some potential causes.

New Contribution

This article presents the first in-depth exploration of how the complexity of public reports might impair consumers' choices among health care providers. To this end, we provide a synthesis of ways in which complex choice sets can affect decision making, as well as the first test of their relative influence on health care choices, in this case, selecting physicians.

Although our findings offer a number of insights, we urge readers to treat them as exploratory rather than definitive, since this is the first foray into a complicated set of behaviors and the experiment on which these findings are based was designed to document the existence of impaired choice rather than to establish which factors most affected decision making. Moreover, we assess consumers' actions against a standard for "good choices" based on a robust use of standardized quality metrics—one plausible notion of fully informed choices but hardly the only reasonable one. Despite these limitations, our findings can be used to illuminate some general strategies for enhancing medical consumerism in the context of public reporting of provider performance. In the Discussion section of the article, we explore some of these implications.

Conceptual Foundation: What Makes Complex Choices More Difficult?

Market-oriented health policy in the United States has been guided by the presumption that choices in medical markets are readily made and can be fully informed (Kling et al., 2012). Under these circumstances, maximizing the number and variety of options for consumers is a reasonable way to assure that at least one alternative closely matches each person's personal preferences and individual health needs. To proponents of this strategy, evidence that consumers fail to achieve this matching indicates that they are insufficiently motivated to seek out information, resulting in a call for higher levels of cost-sharing as an inducement for consumers to pay greater attention (Robinson & Ginsburg, 2009). The case for expanded choice has also been bolstered by evidence that consumers generally seek out multiple options and report themselves to be more satisfied when choice sets are larger (Bundorf & Szrek, 2010).

Some exceptions have captured policy makers' attention. Most noteworthy was the experience of Medicare beneficiaries in the 1980s with private insurance policies that supplemented Medicare's core coverage, often referred to as "Medi-gap" plans. Persistent evidence of marketing abuses and poor consumer judgment (millions of Medicare beneficiaries purchased multiple Medi-gap policies with duplicative coverage) led to federal reforms that standardized coverage, limited the number of variations from which consumers could choose, and funded state-administered programs to assist beneficiaries in their choices. However, given the circumstances, many observers attributed the poor choices to the cognitive deficits associated with aging, especially in light of evidence that bad purchases were most common among the oldest and most cognitively limited Medicare beneficiaries (Hanoch et al., 2009). As a result, policy makers did not generalize from the Medi-gap episode to the design of policies intended for the rest of the American population (Hanoch & Rice, 2006).

Even when focused on the elderly, market reformers paid little attention to the burdens that choice placed on consumers. With the implementation in 2006 of Medicare Part D, which provides prescription drug coverage through another set of private plans, policy makers again favored a strategy of maximizing the number of plans in each market. In most localities, the number of options exceeded 50 (Hanoch & Rice, 2006). In the aftermath, studies concluded that less than half (in some studies, far less than half) of the Part D plans that had been selected actually provided beneficiaries with the best coverage available in the local market, given their existing prescriptions (Domino, Stearns, & Yeh, 2008; Hanoch et al., 2009; McFadden, 2006; Hanoch, Wood, Barnes, Liu, & Rice, 2011). Most of the research on choices under Part D focused on the number of plan options (since this was a point of policy leverage in that program) as a source of choice complexity, though a few studies also examined the number of performance metrics.

Some of these studies also found that the frequency of poor choices increased with the number of plan options (Hanoch et al., 2009; Hanoch et al., 2011) and the complexity of information provided to consumers (Barnes et al., 2012). Similar patterns have been documented for choices among physicians (Kanouse et al., 2012). When consumers have more doctors to choose among or more types of information available on their comparative performance, they are less likely to pick the doctors who objectively perform best (as measured by standardized performance metrics) and more likely to choose doctors whose care is inferior to other available clinicians.

A negative relationship between the quality of decisions and the richness of consumers' choice sets (the number of options and the variety of metrics) runs counter to conventional thinking about rational decision making. But it is entirely consistent with previous research showing that when people must choose among a large number of options in a complex choice task, they adopt heuristic strategies to reduce the amount of information they must consider—strategies that are “boundedly rational” but can result in a suboptimal decision (Begley, 2011; Payne, Bettman, & Johnson, 1993). Even for decisions that have high stakes, such as those involved in medical care, many consumers will make choices in ways that reflect bounded rationality (Schlesinger, 2010; Schneider, 2006; Shaller, 2005; Ubel, 2002).

This behavior poses a challenge to designers of public reports. Festooning reports with a multiplicity of measures and attributes (e.g., for health plans: corporate ownership, accreditations, years operating in locale; for clinicians: age, gender, specialty training, board accreditations, clinical performance, and patient experience) may satisfy the requirements of a variety of stakeholders and address the diversity of consumers' information preferences, but it may also undermine consumers' ability to use them effectively. Based on the broader literature on decision making, we posit four ways in which an abundance of comparative information can affect how well people process, interpret, and use it (Figure 1). Each pathway leads to impaired choice but in strikingly different ways.

Cognitive Overload

First, the amount of information presented in a report may exceed the capacity of boundedly rational consumers to interpret. “Cognitive overload” can undermine decision making in a variety of ways highlighted in the literature. Two seem particularly relevant to clinician

choice: by inducing consumers to satisfice or by encouraging them to filter the information arbitrarily. Some people appear to process information until they hit some threshold (“My brain is full”), then truncate the learning stage of the choice process if it appears that one of the options they have encountered is “good enough” (Hibbard, Slovic, & Jewett, 1997). Unless choices can be ranked along a single dimension and are presented that way to consumers, this sort of stopping threshold can lead them to overlook the best options that are available to them but positioned lower on the list of alternatives.

Other consumers may respond by constraining their choice sets in ways that may help the decision process but sometimes lead to suboptimal decision making. They adopt a “data reduction” strategy, parsing the choice set to reduce the cognitive burden of choice in a manner that roughly matches their personal preferences (Hanoch et al., 2011). In health care, for example, data reduction might involve applying screening criteria, such as looking only at health plans located within 5 miles of home or considering only doctors whose gender matches one’s own.

We refer to such strategies as “filtering” the choice set. Filtering based on just one or two attributes from among a larger number about which the consumer cares poses the risk that some options that are removed from consideration by filtering might actually be superior in meeting the consumer’s preferences if all options and attribute combinations were considered.

Reduced Coherence

A second source of impaired decision making involves the coherence of the comparative performance data—a coherence that may be threatened when public reports incorporate an ever-growing number of measures. More information is not necessarily more difficult to interpret if the additional data clarify distinctions (Glockner & Betsch, 2012). For example, if the additional metrics are strongly correlated with prior measures, their inclusion could make each of the individual measures seem more reliable and trustworthy while also making it easier for consumers to single out the best performers, since each of the measures might seem most relevant to different subsets of consumers.

Nonetheless, there are several ways in which additional measures can make choices harder. If performance metrics are uncorrelated—or even negatively correlated—then consumers must strike trade-offs among attractive attributes. Some consumers find it painful to knowingly give up something that they value, even if it can be justified by gains in other attributes that are valued more (Schwartz et al., 2002). Increasing the domains of measured performance causes the trade-offs to increase multiplicatively, rendering it harder for some consumers to even keep track of the balances that are being struck (Payne et al., 1993).

Introducing additional metrics can also reduce coherence if attributes are measured in ways that make it hard to value trade-offs. How, for example, should a consumer trade off an additional \$100 in monthly premiums against a 5% increase in the number of plan enrollees who have had their appropriate screening exams? How much should one angry, impassioned, and seemingly insightful complaint about the rudeness of office staff be weighed against the frequency of eye exams given to patients with diabetes in a clinician’s practice?

Lost Perspective Due to Cognitive Tunneling

A third source of potentially impaired decision making involves consumers focusing so much on the detailed results of particular measures that they lose a broader perspective. As interactive websites proliferate, consumers increasingly have the ability to “drill down” to examine the components of composite measures, to view the full frequency distribution for scores rather than means, to see time trends in performance scores, or to explore in greater detail the ways in which given measures were constructed. This behavior has some value in terms of engaging consumers in more active learning and increasing their trust in the comparative performance data. Yet it also creates the risk of tunneling: that is, diving deep into particular measures at the expense of others, making it more likely that crucial trade-offs among metrics are overlooked, even though the consumer would have wished (under less demanding circumstances) to have taken them into account.

Emotional Heuristics

Finally, complex choice sets can undermine decision making by heightening the emotional valence of the decision-making process. Choices are generally shaped by “dual-process” models that involve a combination (and integration) of emotion and cognition (Vohs, Baumeister, & Loewenstein, 2007). Under some circumstances, emotions can enhance decision making, as when anxieties cause decision makers to pay closer attention to information that they might otherwise overlook. But there are other circumstances in which emotions can impair judgment or lead people to avoid choosing altogether; several of these circumstances are exacerbated by more complex choice sets (Luce, 1998).

As noted above, an abundance of performance metrics makes it more likely that consumers will need to strike trade-offs among desired attributes or outcomes. People with certain personality types seem to have a particularly hard time with this, often regretting the compromises they feel forced to make (Schwartz et al., 2002). In more extreme cases, regret avoidance can cause people to avoid making choices at all (Anderson, 2003). If forced to choose, these individuals may simply go with the default option if one exists (i.e., passively choose), defer the choice to someone else, or make a random selection, any one of which ensures that they cannot be held responsible by others or themselves for any negative consequences, particularly those that in retrospect might seem to have been misguided trade-offs. In other words, the specter of making a choice that they would later regret impairs their capacity to choose—and, ironically, yields outcomes that they might have wished to avoid.

Studying the Impact of Complexity on Consumers’ Choice of Clinicians

Although increasing the number of options and increasing the number of performance metrics both create more complex choice sets, these features may affect different consumers in different ways for various types of choices. Choices of physicians, for example, are often made with less comprehensive or reliable metrics of performance than is true for health plans, particularly at the level of the individual clinician. Whether this leaves consumers more or less vulnerable to the impact of complexity is an open question.

To begin to explore these implications of complex choice sets for the perceptions and behavior of medical consumers, we constructed a fictitious website for selecting physicians, emulating features found on real-world websites. We describe here this experimental platform, the random assignment of participants to choice contexts of varying complexity (in both senses identified above), our measures of good and impaired decisions, and the specific comparisons used to discern factors that might be influencing the relationship between complexity and choice.

Experimental Context

The Decision Setting—To develop a realistic setting for studying medical consumerism, we created a prototype website called SelectMD that presents comparative quality ratings for a set of fictional physicians. The website is designed with content, format, design, and functionality similar to what participants would encounter on typical “report card” websites.

Participants were presented with performance information on clinicians in their choice set; performance was represented by one to five stars, with three stars representing average performance. They had the choice to begin exploring the website on either a performance overview page—one that presented a single rolled-up CAHPS (Consumer Assessment of Healthcare Providers and Systems) score and (in some experimental arms) a single rolled-up HEDIS (Healthcare Effectiveness Data and Information Set) score—or a page that displayed the four components that were combined to create the rolled-up scores for CAHPS and HEDIS. Virtually all participants chose to start on the overview page; an example of what they observed as they logged onto this page is presented in Figure 2.

Depending on the experimental arm to which they were assigned, some participants were also able to view “comments” volunteered by other patients of the physician under consideration. Participants were able to sort doctors by level of performance on each measure and to “drill down” (aka “probing” the measures) within a given performance category to view component measures. A final feature of the SelectMD website, not visible to participants, was a tracking system monitoring the online use of the website by participants. The system was designed to facilitate analysis of patterns of use by recording every “click” made by participants as well as the time spent on each page.

Participants were asked to use the information to choose a new primary care physician. They could select any subset of clinicians to array on the screen to make choices easier. For verisimilitude, they were allowed to indicate the distance they were willing to travel to a doctor’s office and the type of primary care physician they preferred, though these preliminary selections had no impact on the choice set to which they were exposed.

These efforts to present the choice task in familiar terms seemed successful: 90% of participants rated the website very or somewhat easy to use, and 88% would recommend that their family and friends use a comparable website when choosing a doctor. Before and after visiting the SelectMD website, participants were asked to complete a brief set of questions: beforehand, about their health care experiences, and afterward, about their experiences on the website and decision-making styles.

Experimental Design—To inform their hypothetical choice, all participants could view survey results on patients' experiences with care from each physician. The measures and scores emulated the standard format for reporting survey results from the CAHPS Clinician & Group Surveys (hereafter "CAHPS"), including ratings of the overall care from the clinician, provider–patient communication, timely access to care, and interactions with office staff. All participants could choose among at least a dozen clinicians, who could be sorted by gender and experience. The experiment introduced greater complexity in two ways. The first involved additional forms of quality information: measures of clinical performance drawn from medical records (i.e., HEDIS metrics) and patient anecdotes about their experiences with this doctor (referred to as comments). The HEDIS metrics had explanatory labels with scores presented with the same star ratings used for the survey results. The patient comments displayed in SelectMD were modeled on actual patient anecdotes harvested from real physician rating sites (Martino, Kanouse, Schlesinger, Shaller, & Rybowski, 2012).

Specific comments were assigned to each clinician using a stratified random process. The allocation of comments to physicians followed a matching protocol to ensure that physicians with higher CAHPS ratings (more stars) were assigned more positive comments; however, there was some variance in the emotional valence of comments for all clinicians in the choice set. Specific comments drawn from each tier of emotional valence were randomly assigned to particular clinicians. When comments were available, four to six were presented for each clinician.

The second form of added complexity involved doubling the number of clinicians in the choice set from 12 to 24. Budget limitations prevented us from using a design that fully interacted all of these attributes. The experimental results presented here are based on a design that divided subjects into six equal-sized groups, the first exposed to simplest quality choices (CAHPS measures only, smaller numbers of doctors) and the other five to increasingly complicated choice sets (Figure 3).

Sampling and Participants

Participants were recruited randomly from an existing survey panel of more than 50,000 households developed and maintained by Knowledge Networks, Menlo Park, California (now operating under the GfK corporate label). This panel was constructed using a combination of random digit dialing and address-based sampling, a technique that has been demonstrated to accurately represent the noninstitutionalized U.S. population, including households with unlisted phone numbers, cell phone-only households, and nontelephone households. We restricted eligibility to those who access the Internet via computer, the types of people likely to seek online information about doctors. Of the 1,757 panel members aged 25 to 64 years who were invited to participate, nearly half (48.3%) accepted.

The experiment was conducted in the homes of participants using their Internet connections. This approach creates the potential for disruption. A small number of participants (less than 2%) were interrupted by events that led them to break off their involvement with the experiment for extended periods. We tracked these break-offs and excluded cases in which the break was lengthy.

Measures of Decision Quality

We measured decision quality in two ways. The first assessed whether the selected clinician had been the highest rated on CAHPS (for Arms 1, 3, and 5) or CAHPS plus HEDIS (for Arms 2, 4, and 6). The second assessed whether the participant chose a clinician who was a dominated choice (i.e., there was at least one doctor who scored at least as well on one rolled-up standardized metric and better on the other). Dominance could be assessed only for Arms 2, 4, and 6. Dominated options represent poor choices because the participant could have chosen better no matter what relative value was placed on CAHPS versus HEDIS measures.

To be sure, thoughtful choices among clinicians need not be limited to standardized performance metrics such as CAHPS or HEDIS. We know, for example, that consumers view clinicians' training as a predictor of good care (Hibbard & Jewett, 1996). But in this context, participants had only limited information about clinicians' specialty training, and all the clinicians in their choice set met their preferred criteria (if they chose to make it a precondition). Consumers care about geographic proximity of their doctors' offices, but here again they could set a threshold and all the clinicians in their choice set were putatively located within this distance. Consumers may also care about the sorts of experiential factors that are represented by the patient comments. In this experiment, however, there was no indication that the handful of anecdotes available to them was representative of the clinicians' full patient population; whether this was widely recognized by participants is unclear.

Assessing the Impact of Complexity on Consumers' Choice of Clinician

The core of our analysis involves testing for the effects of each of the causal pathways identified above. Some of these tests are structured into the experimental design; others call for additional measures based on questions asked of participants after they had completed the choice experiment. In the latter case, we assessed their impact by dividing participants into subgroups and comparing across these groups—and, ultimately, by incorporating these group identifiers into a regression model.

Cognitive Overload—A straightforward test of overload is provided by comparing choices made with the same array of performance metrics, but with more clinicians to select among. To do this, we compared the quality of choices between Arms 3 and 5, both of which include CAHPS and comments, and between Arms 4 and 6, both of which include CAHPS, HEDIS, and comments. If participants are satisficing in response to cognitive burdens, one would expect to see poorer choices in Arms 5 and 6, compared to their corresponding versions with fewer clinicians (Arms 3 and 4, respectively).

As noted above, decision makers may respond to cognitive burdens by shrinking the choice set from which they must choose. In the context of the SelectMD website, participants could filter choices by looking only at clinicians of a particular gender or level of experience and selecting their preferred clinician from this restricted choice set. Eighteen percent of participants filtered in this manner. Interestingly, filtering was *not* more common for more complex choice sets, although the impact of filtering did differ (see below).

Coherence—As noted earlier, additional metrics can reduce the coherence of performance assessments in two ways: first, by making the rankings of clinicians inconsistent, and second, by introducing incommensurate ways of interpreting “good” performance. In our experiment, the introduction of HEDIS measures captures the first challenge to coherence; as with actual data on clinicians, the HEDIS scores in the experiment are not correlated with the CAHPS scores (in the literature, some studies find positive correlations between experiential and clinical measures, others negative correlations, and still others no statistical relationship at all). The effects of this reduction in coherence on choice should be evident by three pairwise comparisons where HEDIS measures were added to otherwise comparable choice sets: Arm 1 with 2, Arm 3 with 4, and Arm 5 with 6.

Introducing patient comments is intended to capture the second form of reduced coherence. Not only are anecdotes more laden with emotion—and therefore processed in different ways from the CAHPS or HEDIS ratings—they also cannot be ranked in a straightforward fashion by counting stars, which forces consumers to construct their own way of sensibly interpreting relative performance. Because the comments were assigned so that the emotional valence of the comment reinforces the CAHPS rankings, they may actually add to the first form of coherence for Arms 1, 3, and 5 but may reduce both forms of coherence when HEDIS rankings are also available (Arms 2, 4, and 6).

Although ease of ranking and correlations among measures can make it easier or harder to compare doctors, consumers’ ability to discern coherent patterns in complex arrays of information may also affect decision quality. Some consumers have a greater facility than others at interpreting multidimensional matrices of performance information. This capability is related to other measures of cognitive capacity (first-order correlations of .50–.60; see Bruine de Bruin, Parker & Fischhoff, 2007) but is distinctive enough to merit attention in this context. Drawing on a protocol developed by Payne et al. (1993), participants were asked to imagine that they were assisting a friend in the purchase of a DVD player, using a matrix ranking different brands along three attributes. Each participant was asked to make four different selections, each applying a different decision rule. Just over a third (36.6%) of all respondents correctly identified the preferred choice in at least three of the four test cases: They are identified here as high-skill decision makers.

Cognitive Tunneling—The tendency for consumers to become lost in particular aspects of complex information sets—and therefore to conduct a less balanced assessment of the options among which they are choosing—is likely to be partly influenced by decision styles. Past research has identified a subset of decision makers prone to “excessive maximization”: a tendency to seek the very best option, even if this requires intensive effort (Schwartz et al., 2002). Symptoms of this propensity include agreement with statements like “If I make a choice and it turns out well, I still feel like something of a failure if I find out that the other choice would have turned out better” (a question answered affirmatively by 19% of the participants in this study).

Excessive maximizers may actually make better choices when faced with relatively simple data sets, since they are motivated to get their choices “right.” But this same motivation may lead to the tunneling effect described above when some performance metrics are more

nuanced or difficult to rank. Participants answered questions from an index developed by Schwartz et al. (2002). We combined these additively and identified the top third of the distribution for the combined measure as predisposed to excessive maximization; given the natural breaks in this distribution, this threshold included 37% of respondents.

To test more directly whether consumers were distracted from a more balanced assessment of performance metrics, we also measured whether consumers had examined performance on the components of the summary ranking based on the CAHPS measures. Scores for CAHPS were available in every arm of the experiment; most respondents reported them to be useful for choosing among clinicians. But only 38% of participants took the time to explore the component scores on which the rolled-up CAHPS score was based. We explored whether this behavior varied with the complexity of the choice set and whether the interaction of decision styles and choice set attributes affected decision quality.

Emotional Heuristics—To assess the emotional overtones that might accompany choices made in more complex circumstances, we drew on two indices from the literature. The first, also developed by Schwartz et al. (2002), measures the extent to which participants report being prone to regret over past choices. We anticipate that those with this tendency will be more likely to exhibit the forms of regret avoidance identified above, potentially defaulting to random selection to avoid blameworthy errors. Our goal again was to identify individuals in the upper third of the response distribution; the break points in the distribution led to our labeling 37% of respondents as highly disposed to regret.

Negative emotional valences may also lead to avoiding choices, rather than simply aversion to regret. In contrast to regret avoidance, we anticipate that choice avoidance will induce the largest choice pathologies in the most complex choice sets, since decision makers will be most inclined to distance themselves from responsibility when they can least make sense of their options. Drawing on an index developed by Scott and Bruce (1985), we identified the extent to which participants were avoidant of choice; this index was moderately correlated (.37) with the index of regret avoidance, with sufficient independent variance to distinguish among the two influences. We identified 27% of respondents in this study as highly choice avoidant.

Statistical Methods

Some of the causal pathways assessed here rely on comparisons across experimental arms. Because respondents are randomly assigned to these arms, one can compare across them without controlling for other sociodemographic factors or personality attributes that might be influencing decisions. Other tests, including those based on consumers' actions on the website (filtering and probing) and decision propensities, involve behaviors and attitudes that are not randomly assigned, and thus may be influenced by a variety of other attributes.

To assess the impact of these influences—and their mutual interactions—we therefore estimated a set of logistic regression models, with our measures of good choices (picking the top ranked clinician according to standardized metrics) and bad choices (picking a dominated clinician) as the dependent variables in the regressions. These models were estimated in stages, beginning simply with variables identifying the respondent's

experimental arm, then sequentially introducing decision styles/skills, website behaviors, and interactions between experimental arms and decision propensities.

Findings

We begin by replicating our previously reported findings (Kanouse et al., 2012), documenting that the quality of decision making declines in the face of more complex choices sets. We then examine each of our four causal pathways one-by-one culminating with estimates from our staged logistic regression models to identify the relative importance of each sources of impaired choice.

Declining Decision Quality As Choice Sets Grow More Complex

As can be seen in Figure 4, as one moves from the simplest choice scenario (Arm 1) to the most complex (Arm 6), there is a decline in the proportion of participants who selected clinicians who scored best on standardized performance metrics (CAHPS in Arms 1, 3, and 5; CAHPS + HEDIS in Arms 2, 4, and 6). Each of these pairwise comparisons (Arms 1 with 3; 3 with 5; 2 with 4; 4 with 6) yields a statistically significant difference in decision quality ($p < .05$).

Because dominance can be defined only for choice settings in which performance was ranked along multiple dimensions, this measure can be assessed only for Arms 2, 4, and 6. Again, complexity is associated with suboptimal choice. As one moves from simpler to more complex choice sets, the prevalence of dominated choices increases. In other words, consumers were more likely to choose doctors who scored *less well* on either HEDIS and CAHPS rankings (and no better on the other) than at least one other physician in the choice set. In the most complex choices, half of all selected physician were dominated choices, a prevalence consistent with the frequency of dominated options selected by consumers choosing among health plans (Domino et al., 2008; Hanoch et al., 2011). Both cross-arm comparisons are statistically significant ($p < .05$).

Causal Pathways Linking Complexity to Decision Quality

Cognitive Overload—Our first-order test for cognitive overload involves comparing Arm 3 with Arm 5 and Arm 4 with Arm 6. As noted above, the frequency with which consumers selected the top-ranked doctors was significantly lower and the frequency with which they selected dominated doctors significantly higher for the arms offering more clinicians to select among, though only the difference between Arms 3 and 5 was statistically significant ($p < .01$).

Consumers who look for ways to cope with cognitive overload may be inclined to filter their options based on some easily applied attribute. Experimental participants could filter based on either years of experience or the gender of the clinicians available to them; filtering had clear consequences for the quality of their decisions but in a complicated manner (Figure 5). In the simplest choice set (Arm 1), filtering was entirely unrelated to decision quality. By contrast, in Arms 2 to 5, those who filtered were *far* less likely to select the top-ranked physician and more likely to select a dominated option; all these differences within the four arms are statistically significant ($p < .05$).

Arm 6 appears to be a bit of an anomaly. Dominated choices are elevated to a statistically significant extent for those who filtered. But participants are only slightly less likely to have selected the top-ranked clinician in their choice set.

Coherence—We aimed to assess the impact of two distinct forms of reduced coherence. The first, involving correlations among performance metrics, entails comparing Arm 1 with Arm 2, Arm 3 with Arm 4, and Arm 5 with Arm 6. None of these comparisons are associated with statistically significant differences in choice of the top-ranked clinician. This suggests that coherence in the sense of uncorrelated aspects of performance does not significantly impair decision quality.

By contrast, coherence in the form of incommensurable forms of assessment introduced in the form of patient comment seems to affect decision quality more profoundly: as observed above, participants in Arms 3 and 4 were significantly less likely to have selected the top-ranked clinicians than were participants in Arms 1 and 2 (Figure 4). The decline between Arms 2 and 4 is particularly pronounced, suggesting that the effects of patient comments may interact with the addition of the HEDIS measures ($p < .01$).

We hypothesized that the consumers most able to interpret matrices of performance data might better cope with the potential incoherence of complex performance metrics. When we compared more and less skilled consumers, the differences across arms were not generally significant. There was one exception however: Arm 6. Faced with this most complex choice set, skilled consumers were more likely to select the top ranked physicians (48.9% vs. 31.7%; $p < .05$) and less likely to select a dominated clinician (37.8% vs. 55.5%; $p < .05$).

Cognitive Tunneling—We deployed two measures that are potentially related to cognitive tunneling. The first was whether a participant probed (“drilled down”) to see component measures of CAHPS. Those facing more complex choice sets were less likely to have done so, but these differences were modest in magnitude and not statistically significant. Nor were the differences consistently associated with differences in the quality of decision making.

Our second measure was more consistently related to decision quality (Figure 6). We hypothesized that consumers prone to excessive maximization could get caught up in the nuances of specific measures, undermining their ability to make good decisions using multiple metrics. But this decision style did not impair selection when choice sets included only standardized metrics (Arms 1 and 2). In fact, high maximizers were *more* likely to select the top-ranked doctor and less likely to choose a dominated option in these circumstances, though the latter difference was not statistically significant. But the striking contrast came when comments were introduced (Arms 3 and 4). Maximizers became far less able to select the top-ranked doctors and far more prone to selecting dominated options ($p < .05$). These effects, however, appeared to be moderate in the choice sets that introduced more clinicians.

Emotional Heuristics—Neither regret avoidance nor choice avoidance was consistently related to the quality of decision making. But both evidenced the pattern identified above for maximizers, albeit in a more uneven fashion. For the simpler choice sets (Arms 1 and 2),

each of these emotional factors was associated with somewhat higher quality decisions, a pattern that was reversed with more complex choice sets (these interactions are explored in greater detail in the regression results presented below).

Parsing Influences: Findings From Regression Models

To sort out the interactions and relative influences among these different pathways, we estimated a set of regression models (Table 1). In the baseline models (the leftmost column of results in the table), the only variables were dichotomous variables identifying the arm of the experiment; results are presented here in comparison to Arm 1 (Arm 2 for dominated choice). These results replicated the findings for the bivariate comparisons: statistically significant differences for arms that presented more complex choice sets, except when additional standardized metrics were the only source of increased complexity (e.g., Arm 1 vs. Arm 2). The estimated effects of complexity on dominated choice were particularly large: Dominated choices were more than four times as likely for Arm 6 as for Arm 2.

Incorporating decision style and skills into the models did little to alter these patterns (second column of results from the left), though those skilled in interpreting matrices of performance information were less prone to selecting dominated options. Nor did the addition of variables for filtering or probing alter these cross-arm comparisons (third column of results), even though both of the latter behaviors were consistently linked to decision quality. (Filtering was associated with poorer quality decisions, probing with better choices, at least when measured by dominated options.)

However, in models that interact decision style with each experimental arm (far right column in Table 1), a strikingly different pattern emerged. (The models in which these full interaction results were estimated are reported in Technical Appendix A; see appendix, available in the online version of this article at <http://mcr.sagepub.com/supplemental>.) As these results demonstrate, the decline in decision quality that was associated with the introduction of patient comments in Arms 3 and 4 was entirely attributable to poorer quality decisions among consumers prone to excessive maximization. (Though the frequency of dominated choices remained elevated for Arm 4 even after accounting for these interactions, this was no longer a statistically significant difference.) Strikingly, a propensity toward excessive maximization was *not* associated with worse decisions for simpler choice sets. Indeed, quite the opposite was the case: This subset of consumers appeared to make consistently better choices for Arms 1 and 2, though not all of these differences were statistically significant.

Equally noteworthy, maximization did not seem to account for the lower decision quality in Arms 5 and 6, even though these choice sets also included patient comments. Some of the poor quality choices in these most complex arms appeared to be attributable to choice avoidance (particularly for dominated choices in Arm 6). But even after accounting for this decision style, the frequency of poor quality decisions remained elevated in Arms 5 and 6, suggesting that some additional factors beyond those identified here were coming into play. Here again, decision styles associated with lower quality choices in complex contexts were associated with *better* choices when the choice set was simpler. For example, respondents

prone to choice avoidance were twice as likely as other consumers to choose the highest rated clinician in Arm 1.

Finally, a more subtle pattern was evident in these findings, though the statistical significance is only borderline. Respondents who were more sensitive to emotional aspects of choice (those who were high on regret or choice avoidance) appeared to have a significantly easier time selecting the best clinician when presented with a single standardized metric (CAHPS) rather than two (CAHPS plus HEDIS). Unfortunately, the design of the experiment does not make it possible to discern whether this reflects a differential response to the two types of standardized measures or to a multiplicity of standardized measures and the trade-offs that ensue.

Discussion and Conclusion

Greater complexity in reports on the performance of health care providers affects the ways in which consumers assess their options and make their decisions. Although richer choice sets have the capacity to engage consumers more fully (Bundorf & Szrek, 2010; Kanouse et al., 2012), they also clearly tax the abilities of consumers to make good choices—particularly among those with more limited decision skills and particular decision styles. These are not small differences: Those most at risk made “questionable choices” 4 times more often than did other consumers. Other factors (e.g., regret avoidance) identified in the broader literature on decision making seem less consequential for health care choices, at least those made under the circumstances represented by this study.

Methodological Limitations

It is important, of course, to view these findings in light of certain methodological considerations. As we emphasized above, with few previous studies examining the decision processes involved in choice of clinicians, we have only begun to explore the ways in which complexity may matter. Future research will need to expand the scope of inquiry to include the influence of a wider array of performance metrics (e.g., cost) and provider attributes (e.g., place of training) as well as additional options for filtering the choice set. The HEDIS measures included on the SelectMD website (two measuring treatment of chronic conditions, two measuring preventive care) were also not meant to exhaust all possible metrics for clinical performance. There is more to be learned by exploring different sets of clinical and experiential measures, including those related to patient safety, shared decision making, and decision support by clinicians—as well as different means of presenting both types of metrics in an integrated fashion. It will also be important to extend this exploration to other health care-related choices, such as health plans, specialists, and other types of health services.

Moreover, our findings suggest that some of the impact of complexity involves factors beyond the four pathways identified here. These findings raise puzzling questions that require additional study about how consumers cope with the most complex choice sets.

Perhaps most consequentially, it is essential to extend the study of how consumers select clinicians from hypothetical choices, like those posed in our experiment, to real-world

contexts. This transition has already occurred for the study of health plans (Faber et al., 2009), though none of the studies of real plan choice have examined the pathways that lead to these decisions. There has been some preliminary research of this sort (Martino, Kanouse, Elliott, Teleki, & Hays, 2012); it will be important to further develop a robust portfolio of research on actual choices of clinicians.

Implications for Enhanced Reporting

Despite these limitations—and the provisional nature of our findings—it may be useful to consider some implications for reporting health care quality to the public. First, it is clear that complexity matters. The quality of choice erodes dramatically (in the sense that information contained in quantitative metrics is neglected) as choice sets incorporate more options and more performance metrics, even though the information on SelectMD was far simpler than on websites that also include cost measures, safety metrics, and assessments of electronic access. The push to incorporate an ever larger array of performance measures comes at a real cost. Enhancing choice may call for making some metrics *less* available, visible only to consumers who are sufficiently motivated to actively seek them out on a website.

Concerns about bounded rationality are often addressed by calling for “roll-up” scores that combine diverse measures into a single metric, for example, a single CAHPS score or a single HEDIS score. SelectMD incorporated exactly these sorts of aggregated scores. The patterns of website use that we observe make us cautious about its wider application. Relatively few consumers probed below the roll-up scores to their component measures. Undoubtedly, many chose not to drill down to the components because they saw little value in doing so, but in the absence of at least exploring the component measures and examining their intercorrelation in this choice set, consumers lost some vital information about potential trade-offs among the different measures combined in the roll-ups.

Second, certain types of consumers are more susceptible than others to the challenges of complexity. Some of these susceptibilities relate to decision skills, others to decision styles. Interactive websites could collect information by asking consumers about their decision skills or propensities and then use this information to alter how metrics are presented to each consumer. When these conditional displays simplify the initial presentation of metrics to facilitate choice, the potential for drilling-down could be accompanied by warnings designed to reduce the propensity for some consumers to get “stuck” in the details of a single domain of performance.

Third, incorporating patient comments into public reporting clearly displaces consumers’ attention from standardized performance metrics, leading to choices that are suboptimal on those measures. But narratives do engage patients, in part by conveying information about clinicians that is richer in meaning and emotional nuance (Lagu & Lindenauer, 2010). Report designers must find ways to incorporate these comments without undermining the use of standardized metrics that are more representative than anecdotes of the range of patients’ experiences with each clinician.

Fourth, giving consumers the ability to filter choice sets on interactive websites presents a double-edged sword. On one hand, use of these filters helps relieve cognitive burden by reducing the size of a choice set to one that seems manageable. On the other, in so doing, consumers often exclude the best clinicians. It remains unclear how best to alert consumers to this risk, but the need to do so is apparent.

Finally, it is essential to recognize that there is no single magic bullet for remedying the problems that currently beset public reporting (Sinaiko et al., 2012). Consumer choices are shaped by a multiplicity of factors; we identified half a dozen influences, each requiring a distinctive response. Even more challenging, some of the very factors that inhibited good choices in complex settings actually led to better decision making when the choice set was simpler. Under these circumstances, it is essential to develop a portfolio of strategies that can assist consumers and deploy those strategies in ways that fit consumers' circumstances.

Conclusion

A better understanding of how people make sense of complex information is paramount for designing more thoughtful policies to improve health system performance. So too is a more comprehensive perspective on consumer empowerment. The policy challenge is not simply to transform Americans into better or more motivated medical consumers but also to assist consumers with making sense of choices that will inevitably be challenging, no matter how engaged they are in these choices. Moreover, performance reporting is not simply about choice; it is also intended to educate people and enhance their capacity to engage in shared decision making and to voice their concerns if their health care seems to have gone awry.

Given these multiple goals, policy makers must approach performance reporting in a more nuanced manner—and with greater attention to the variety of strategies that are needed to adequately support Americans as they negotiate their way through the health care system. Recent reforms under the Patient Protection and Affordable Care Act of 2010 are steps in this direction, melding requirements that states publicly report on the performance of health plans with commitments to provide “insurance navigators” and other forms of assistance to help consumers sort through their options (Day & Nadash, 2012; Grob et al., 2013). Nonetheless, more effectively supporting and empowering Americans to render judgments and make considered choices in health care settings is a challenging task that will require a serious commitment of resources and diligent attention from public officials.

Acknowledgments

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Support for the research was provided by two cooperative agreements (UI18HS016980 and 1U18HS016978) from Agency for Healthcare Research and Quality to RAND and Yale University, respectively.

References

- Anderson C. The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin*. 2003; 129:139–166. [PubMed: 12555797]

- Barnes AJ, Hanoch Y, Wood S, Liu PJ, Rice T. One fish, two fish, red fish, blue fish: Effects of price frames, brand names, and choice set size on Medicare Part D insurance plan decisions. *Medical Care Research and Review*. 2012; 69:460–473. [PubMed: 22311955]
- Begley, S. I can't think!. *Newsweek*. 2011 Feb 27. Retrieved from <http://www.newsweek.com/2011/02/27/i-can-t-think.print.html>
- Bruine de Bruin W, Parker A, Fischhoff B. Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*. 2007; 92:938–956. [PubMed: 17484614]
- Bundorf KM, Szrek H. Choice set size and decision making: The case of Medicare Part D prescription drug plans. *Medical Decision Making*. 2010; 30:582–593. [PubMed: 20228281]
- Day R, Nadash P. New state insurance exchanges should follow the example of Massachusetts by simplifying choices among health plans. *Health Affairs*. 2012; 31:982–989. [PubMed: 22566437]
- Domino ME, Stearns SC, Yeh WS. Why using current medications to select a Medicare Part D plan may lead to higher out-of-pocket payments. *Medical Care Research and Review*. 2008; 65:114–126. [PubMed: 17942795]
- Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: How do consumers use quality-of-care information? A systematic review. *Medical Care*. 2009; 47:1–9. [PubMed: 19106724]
- French, M., Holmes-Rovner, M., Kanouse, D., Parente, S., Shaller, D., Sofaer, S. Getting tools used: Lessons for health care from successful consumer decision aids. Washington, DC: Center for Advancing Health; 2009.
- Fung CH, Lim Y, Mattke S, Damberg C, Shekelle PG. Systematic review: The evidence that publishing patient care performance data improves quality of care. *Annals of Internal Medicine*. 2008; 48:111–123.
- Glockner A, Betsch T. Decisions beyond boundaries: When more information is processed faster than less. *Acta Psychologica*. 2012; 139:532–542. [PubMed: 22381940]
- Grob R, Schlesinger M, Davis S, Cohen D, Lapps J. The Affordable Care Act's plan for consumer assistance with insurance moves states forward but remains a work in progress. *Health Affairs*. 2013; 32:347–356. [PubMed: 23381528]
- Hanoch Y, Rice T. Can limiting choice increase social welfare? The elderly and health insurance. *Milbank Quarterly*. 2006; 84:37–73. [PubMed: 16529568]
- Hanoch Y, Rice T, Cummings J, Wood S. How much choice is too much? The case of the medicare prescription drug benefit. *Health Services Research*. 2009; 44:1157–1168. [PubMed: 19486180]
- Hanoch Y, Wood S, Barnes A, Liu PJ, Rice T. Choosing the right Medicare prescription drug plan: The effect of age, strategy selection, and choice set size. *Health Psychology*. 2011; 30:719–727. [PubMed: 21604880]
- Hibbard JH, Greene J, Sofaer S, Firminger K, Hirsh J. An experiment shows that a well-designed report on costs and quality can help consumers choose high-value health care. *Health Affairs*. 2012; 31:560–568. [PubMed: 22392666]
- Hibbard JH, Jewett JJ. What type of quality information do consumers want in a health care report card? *Medical Care Research and Review*. 1996; 53:28–47. [PubMed: 10156433]
- Hibbard JH, Slovic P, Jewett JJ. Informing consumer decisions in health care: implications from decision-making research. *Milbank Quarterly*. 1997; 75:395–414. [PubMed: 9290635]
- Iyengar, SS., Jiang, W., Huberman, G. How much choice is too much? Determinants of individual contributions in 401k retirement plans. In: Mitchell, OS., Utkus, S., editors. *Pension design and structure: New lessons from behavioral finance*. Oxford, England: Oxford University Press; 2004. p. 83-97.
- Kanouse, DE., Schlesinger, M., Shaller, D., Martino, S., Rybowski, L. The impact of patient narrative comments on consumer use of physician performance reports. Paper presented at the 2012 Agency for Healthcare Research and Quality annual conference; Bethesda, MD. 2012 Sep.
- Kling JR, Mullainathan S, Shafir E, Vermeulen LC, Wrobel MV. Comparison friction: Experimental evidence from Medicare drug plans. *Quarterly Journal of Economics*. 2012; 127:199–235. [PubMed: 22454838]

- Lagu T, Lindenauer PK. Putting the public back in public reporting of health care quality. *Journal of American Medical Association*. 2010; 304:1711–1712.
- Luce MF. Choosing to avoid: Coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research*. 1998; 24:409–433.
- Martino S, Kanouse D, Elliott M, Teleki S, Hays R. A field experiment on the impact of physician-level performance data on consumers' choice of physician. *Medical Care*. 2012; 50:S65–S73. [PubMed: 23064279]
- Martino, SC., Kanouse, DE., Schlesinger, M., Shaller, D., Rybowski, L. Consumers' appraisal of anecdotal accounts of patient experience. 2012 Sep. Presented at the annual meeting of the Agency for Healthcare Research and Quality
- McCormack LA, Schnaier JA, Lee AJ, Garfinkel SA. Medicare beneficiary counseling programs: What are they and do they work? *Health Care Financing Review*. 1996; 18:127–140. [PubMed: 10165027]
- McFadden D. Free markets and fettered consumers. *American Economic Review*. 2006; 96:3–29.
- Patient Protection and Affordable Care Act of 2010. 2010 Pub. L. No. 111-148.
- Payne, JW., Bettman, JR., Johnson, EJ. *The adaptive decision maker*. New York, NY: Cambridge University Press; 1993.
- Robinson JC, Ginsburg PB. Consumer-driven health care: promise and performance. *Health Affairs*. 2009; 28:w272–w281. [PubMed: 19174389]
- Schlesinger M. Choice cuts: Parsing policymakers' pursuit of patient empowerment from an individual perspective. *Health Economics, Policy and Law*. 2010; 5:365–387.
- Schneider C. After autonomy. *Wake Forest Law Review*. 2006; 41:411–444.
- Schwartz B, Ward A, Monterosso J, Lyubomirsky S, White K, Lehman DR. Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology*. 2002; 83:1178–1197. [PubMed: 12416921]
- Scott SG, Bruce RA. Decision-making style: The development and assessment of a new measure. *Educational and Psychological Measurement*. 1985; 55:818–831.
- Shaller, D. *Consumers in health care: The burden of choice*. Oakland: California Healthcare Foundation; 2005.
- Sinaiko AD, Eastman D, Rosenthal MB. How report cards on physicians, physician groups, and hospitals can have greater impact on consumer choices. *Health Affairs*. 2012; 31:602–611. [PubMed: 22392672]
- Sinaiko AD, Hirth RA. Consumers, health insurance and dominated choices. *Journal of Health Economics*. 2011; 30:450–457. [PubMed: 21300414]
- Ubel PA. Is information always a good thing? Helping patients make “good” decisions. *Medical Care*. 2002; 40(9 Suppl):V39–V44. [PubMed: 12226584]
- Vohs, KD, Baumeister, RF, Loewenstein, G., editors. *Do emotions help or hurt decision making?*. New York, NY: Russell Sage Foundation; 2007.
- Werner RA, Asch DA. The unintended consequences of publicly reporting quality information. *Journal of the American Medical Association*. 2005; 293:1239–1244. [PubMed: 15755946]

General Pathway	Particular Form of Response
Cognitive Overload	<ol style="list-style-type: none">1. Abundant information induces incomplete processing → satisficing2. Abundant information leads to deployment of data reduction strategies
Reduced Coherence	<ol style="list-style-type: none">1. Inconsistent ranking across multiple dimensions induces indecision/confusion2. Incommensurate metrics across dimensions leads consumers to ignore one or more dimensions entirely
Loss of Perspective and Balance	<ol style="list-style-type: none">1. “Deep dive” into data leads to tunneling: Losing forest for the trees
Emotional Heuristics	<ol style="list-style-type: none">1. Regret avoidance induces attention to emotionally resonant measures2. Avoiding choice leads to reduced agency and random selection

Figure 1.
Possible choice pathways linking complex choice sets to impaired consumer choice.

Figure 2.
 Screenshot of performance overview page on SelectMD website: Arm 4.

ARM	PERFORMANCE METRICS	DOCTORS
1	CAHPS	12
2	CAHPS + HEDIS	12
3	CAHPS + COMMENTS	12
4	CAHPS + HEDIS + COMMENTS	12
5	CAHPS + COMMENTS	24
6	CAHPS + HEDIS + COMMENTS	24

Figure 3.

Design for Select MD experiment.

Note: CAHPS = Consumer Assessment of Healthcare Providers and Systems; HEDIS = Healthcare Effectiveness Data and Information Set.

Figure 4.
Complexity of choice sets and quality of decision making.

Figure 5.
Filtering choice sets and quality of decision making.

Figure 6.
Excessively maximizing decision makers: Complexity and quality of decision making.

Table 1

Explaining the Impact of Complexity on Quality of Decision Making: Regression-Adjusted Differences Among Experimental Arms.

	Part 1							
	Baseline modes		+Decision measures		+Selection behaviors		+Interaction terms	
	Odds ratio	Confidence Interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval
Selecting the best doctor								
(Comparison: Arm 1)								
Arm 2	1.02	[0.61, 1.68]	0.94	[0.56, 1.59]	0.85	[0.51, 1.44]	1.05	[0.50, 2.32]
Arm 3	0.62	[0.38, 0.99]	0.62	[0.38, 1.00]	0.55	[0.33, 0.89]	0.85	[0.40, 1.84]
Arm 4	0.49	[0.30, 0.80]	0.49	[0.30, 0.81]	0.46	[0.28, 0.77]	1.10	[0.51, 2.38]
Arm 5	0.34	[0.21, 0.55]	0.32	[0.20, 0.53]	0.30	[0.18, 0.49]	0.42	[0.20, 0.90]
Arm 6	0.37	[0.23, 0.61]	0.37	[0.23, 0.61]	0.34	[0.21, 0.57]	0.75	[0.35, 1.62]
High maximizer			0.80	[0.60, 1.08]	0.79	[0.58, 1.06]	–	–
High regret			1.02	[0.76, 1.39]	1.03	[0.76, 1.40]	–	–
High avoidance			1.16	[0.83, 1.61]	1.19	[0.85, 1.66]	–	–
High selection skills			1.23	[0.92, 1.65]	1.24	[0.92, 1.67]	1.22	[0.90, 1.65]
Filtered data set					0.42	[0.28, 0.64]	0.42	[0.28, 0.64]
Probed CAHPS components					0.99	[0.74, 1.32]	1.02	[0.76, 1.38]
Interaction terms								
High maximizer								
Arm 1							1.12	[0.52, 2.43]
Arm 2							2.35	[1.02, 5.42]
Arm 3							0.38	[0.19, 0.79]
Arm 4							0.44	[0.20, 0.95]
Arm 5							0.92	[0.44, 1.92]
Arm 6							0.67	[0.32, 1.39]
High regret								
Arm 1							1.42	[0.64, 3.12]
Arm 2							0.69	[0.30, 1.56]
Arm 3							1.93	[0.92, 4.03]
Arm 4							0.69	[0.32, 1.49]

Part 1									
	Baseline modes		+Decision measures		+Selection behaviors		+Interaction terms		
	Odds ratio	Confidence Interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval	
Arm 5							1.14	[0.52, 2.53]	
Arm 6							0.63	[0.29, 1.36]	
High avoidance									
Arm 1							1.99	[0.87, 4.56]	
Arm 2							1.26	[0.52, 3.07]	
Arm 3							1.47	[0.67, 3.21]	
Arm 4							0.85	[0.34, 2.13]	
Arm 5							1.24	[0.50, 3.05]	
Arm 6							0.80	[0.36, 1.78]	
Part 2									
	Baseline modes		+Decision measures		+Selection behaviors		+Interaction terms		
	Odds ratio	Confidence Interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval	
Selecting dominated option									
(Comparison: Arm 2)									
Arm 4	2.79	[1.57, 4.94]	2.80	[1.56, 5.02]	2.59	[1.42, 4.74]	1.41	[0.61, 3.32]	
Arm 6	4.68	[2.67, 8.22]	4.36	[2.46, 7.72]	4.27	[2.37, 7.69]	2.35	[1.03, 5.39]	
High maximizer			1.24	[0.79, 1.95]	1.30	[0.81, 2.06]	–	–	
High regret			1.03	[0.65, 1.63]	1.00	[0.62, 1.61]	–	–	
High avoidance			1.31	[0.79, 2.16]	1.36	[0.81, 2.29]	–	–	
High selection skills			0.63	[0.40, 1.00]	0.63	[0.39, 1.00]	0.61	[0.38, 0.99]	
Filtered data set					3.28	[1.84, 5.83]	3.17	[1.77, 5.66]	
Probed CAHPS components					0.56	[0.35, 0.90]	0.55	[0.34, 0.89]	
Interaction terms									
High maximizer									
Arm 2							0.64	[0.21, 1.91]	
Arm 4							1.92	[0.89, 4.12]	
Arm 6							1.34	[0.65, 2.78]	
High regret									
Arm 2							0.81	[0.28, 2.39]	

Part 2

	Baseline modes		+Decision measures		+Selection behaviors		+Interaction terms	
	Odds ratio	Confidence Interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval	Odds ratio	Confidence interval
Arm 4							1.27	[0.58, 2.78]
Arm 6							0.96	[0.45, 2.06]
Choice avoidance								
Arm 2							0.87	[0.27, 2.86]
Arm 4							0.95	[0.37, 2.39]
Arm 6							2.34	[1.05, 5.19]

Note: CAHPS = Consumer Assessment of Healthcare Providers and Systems.