

## **Attachment E SelectMD 2.0 Consumer Choice Experiment**

### **Construction of Patient Reviews**

In the SelectMD 2.0 study, participants in seven of the eight experimental arms (all but Arm 1) will be given the opportunity to read “Patient Reviews” in the form of patients’ anecdotes about their experiences with a physician. These patient reviews will be provided along with information about service quality (CAHPS data on patient experiences), how closely a doctor provides the most effective treatments and preventive care (modeled after HEDIS measures), and data on practices that promote patient safety. To construct patient reviews (anecdotal comments) that seem realistic to research participants and differ systematically in their emotional valence (so that they can be assigned to physicians in a controlled manner), we conducted the formative work described below. This formative work was conducted as part of our previous health care consumer choice study, SelectMD 1.0.

First, we gathered a set of existing patient reviews (hereafter referred to as “real reviews”) from the website RateMDs.com. The real reviews were from patients of physicians at practices in four states: Georgia, Missouri, New Jersey, and Oregon.

Using the real reviews as models, research staff constructed 160 patient reviews (hereafter referred to as “made-up reviews”) that addressed similar aspects of physician performance as those addressed in the real reviews (accessibility, communication, caring about patients, staff quality) but that were also designed to fit into one of four categories of affective tone: strongly negative, mildly negative, mildly positive, and strongly positive. We constructed each statement with a specific affective tone by combining largely emotionally neutral words with either adjectives or adverbs that convey clear emotional tone.<sup>1</sup> Statements written to be strongly negative (positive) include two adjectives/adverbs that carry a strong negative (positive) tone. Statements written to be mildly negative (positive) include only a single emotionally charged adjective/adverb with a moderate tone. For example, “distressed” and “useless” were rated as strongly negative words, “excuse” and “nuisance” as mildly negative, “trust” and “consoled” as mildly positive, and “capable” and “outstanding” as strongly positive.

We subjected our 160 made-up reviews, along with the 121 real reviews, to three rounds of pilot testing. Participants in these pilot tests (N = 8-9 per round) either rated the informativeness and affective tone of the reviews or judged the likelihood that the review was real. Participants who rated the informative and affective tone of the reviews answered the following two questions: (a) “How much does this comment tell you about what it would be like to visit this doctor’s office (1 = tells me nothing at all to 5 = tells me a great deal)?” and (b) “How do you think the person who wrote this comment felt about their experience at this doctor’s office when they made this comment (1 = strongly negative to 5 = strongly positive). Participants who judged the perceived authenticity of the reviews were told, “some of these comments are from real people and reflect real experiences, while other comments have been made up to sound like they come from real people,” and were asked to judge the likelihood that each patient review was real or made-up, using the following 5-point scale: 1 = definitely a made-up comment, 2 = probably a made-up comment, 3 = can’t tell whether real or made-up, 4 = probably a real comment, 5 = definitely a real comment. We asked participants to rate the real as well as the made-up reviews so that we could use the data to attempt to equalize the two

types of reviews on these three dimensions (i.e., affective tone, informativeness, and authenticity). Toward that end, we engaged in a process of revision and deletion of made-up patient reviews between rounds of pilot testing and analysis. Across all rounds of testing, each individual review received between 8 and 19 ratings on each of the three dimensions. Based on these ratings, we assigned each patient review (real and made-up) a mean score on perceived affective tone, informativeness, and authenticity.

Ultimately, we selected from our pool of made-up reviews ones that met several criteria. First, the perceived affective tone of the review had to fit within one of four prescribed bands defined by the rating scale for perceived valence (strongly negative patient reviews had to have an average rating of 1.00-1.49 on the perceived valence scale; mildly negative patient reviews, 1.50-2.49; mildly positive patient reviews, 3.50-4.49; and strongly positive patient reviews, 4.50-5.00). Second, within each category of affective tone, the average length, perceived informativeness, and perceived authenticity of made-up reviews had to match as closely as possible that of real reviews. Finally, the patient reviews selected from each category of affective tone had to collectively represent a diverse mix in terms of the number and type of topics on which they focused.

Our final set of made-up reviews includes 35 that were judged to be strongly negative, 32 that were judged to be mildly negative, 40 that were judged to be mildly positive, and 22 that were judged to be strongly positive. The table on the next page provides descriptive data on the length, perceived informativeness, and perceived authenticity of these 129 anecdotes, by category of affective tone. The table also provides comparable data on the real reviews that were gathered from RateMDs.com.

*Comparison of Real and Made-up Patient Reviews on Length, Perceived Informativeness, and Authenticity by Category of Affective Tone*

Affective Tone	Type of Review	Number of Reviews	Informativeness		Authenticity		Length	
			Mean	SD	Mean	SD	Mean	SD
Strongly negative	Real	17	3.42	0.85	2.99	0.87	55.1	43.2
	Made-up	35	3.66	0.46	2.82	0.49	42.4	10.8
Mildly negative	Real	13	3.03	0.53	2.88	0.76	38.3	37.1
	Made-up	32	3.29	0.50	3.26	0.52	36.3	14.0
Mildly positive	Real	35	2.87	0.65	3.85	0.54	21.4	24.4
	Made-up	40	3.38	0.42	3.72	0.39	30.2	9.5
Strongly positive	Real	56	3.62	0.61	3.75	0.42	45.2	28.9
	Made-up	22	3.86	0.29	3.52	0.70	34.6	13.6

Note:

1. The emotional valence of these adjectives and adverbs was determined by a list constructed by the ANEW research team (Bradley & Lang, 1999), who assessed the emotional content of over 1,000 words in the English language using standardized methods and assigned each word a value for valence (position on a scale ranging from extremely positive to extremely negative).

Reference:

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW). Gainesville, FL. The NIMH Center for the Study of Emotion and Attention, University of Florida.