

Intervention

Testing whether decision aids introduce cognitive biases: Results of a randomized trial<sup>☆</sup>

Peter A. Ubel<sup>a,b,c</sup>

information [7]. Thus, people who learn about the benefits of tamoxifen first will remember the risks of tamoxifen better, and make their judgments accordingly.

In testing DAs, researchers rarely look for such biases. Instead, they typically evaluate DAs by assessing whether the DAs increase patients' knowledge, reduce decisional conflict, increase decision satisfaction and change people's decisions [8]. All these outcomes are important measures of DA quality. But none of these measures are designed to identify cognitive biases [9]. For instance, the way most studies of DAs are designed make it impossible for us to even look for order effects, since the studies typically do not vary the order by which the DAs give patients information.

In testing *The Guide to Decide*, we took two relatively unusual steps. First, we broke our DA into parts that we varied, randomly, across participants. For example, we varied the order with which women received information about the benefits and risks of tamoxifen. By randomizing the order of this information, we could test whether the order influenced people's judgments.

But if we discover order effects through this method, we are potentially left in a quandary: women making this decision need information about both the risks and benefits of tamoxifen, and we have to give them that information in some order. If order effects exist, what can we do in response?

Our second unusual step was to address this problem by varying several other elements of our DA across participants, elements we chose because we thought they could potentially eliminate biases like order effects.

In this manuscript, we present results from a randomized trial of *The Guide to Decide*, exploring whether our DA introduced order effects, and whether a specific element of our DA – presentation of contextual risk information – eliminated such order effects. We present these results to illustrate how researchers can identify whether DAs introduce cognitive biases and to demonstrate a technique for testing whether specific interventions reduce or eliminate these same biases.

## 1. Methods

### 1.1. Participants

We recruited women whose 5-year risk of developing breast cancer was  $\geq 1.66\%$  (estimated by the Gail model) [10]. The baseline risk of 1.66% was chosen based on the minimum risk level used in the NSABP P-1 study [6]. Participants were recruited from two regional health care organizations (in Detroit, MI and Seattle, WA). We used electronic medical data from each organization to determine a preliminary estimate of women's 5-year risk of breast cancer, and whether women had contraindications to tamoxifen (i.e. renal or liver disease, history of pulmonary embolism or deep vein thrombosis). We then mailed eligible women an invitation letter providing a description of the study and the website address for *The Guide to Decide*. Interested women logged onto the site and were further screened for eligibility. Only women aged 40–74 (the ages for which tamoxifen prophylaxis is approved) whose 5-year risk was  $\geq 1.66\%$  were eligible to participate. Other eligibility criteria included no prior history of breast cancer or taking tamoxifen and no participation in clinical intervention studies such as the STAR trial. (The results of the STAR trial had not been released at the time of our study.) Women who had a terminal illness or who were pregnant or nursing were ineligible.

### 1.2. Design of decision aid

Eligible participants were randomized to view 1 of 16 versions of the DA. Each version contained the same basic information, but

differed in the presentation of the key numerical risk/benefit information, with five experimental factors we varied randomly across participants in a fractional factorial design [11].

We chose two of the experimental factors with the specific intent of determining whether our DA created potential cognitive biases. We have already mentioned one of these two factors—the order of risk/benefit information. As discussed above, the DA differed in part by whether women received information about the risks of tamoxifen before its benefits, or vice versa. We hypothesized that participants who received risk information last would be more knowledgeable about the risks of tamoxifen and therefore more worried about them, since such information would be fresher in their minds.

The second potential bias we studied involved the manipulation of risk denominators, to test for denominator biases. Prior research has shown that some people are more worried about a 40 in 1000 risk than a 4 in 100 risk, even though such risks are mathematically identical [12]. To test for this bias, we randomized participants to receive descriptions of risk with either a denominator of 100 or 1000. We discuss denominator effects in more detail elsewhere [13].

We also created three other experimental factors, all of which we predicted could potentially reduce cognitive biases. We have written previously about how two other factors influenced women's knowledge—the use of pictographs to provide a visual display of risk information, and the presentation of information on incremental risk to help people understand the change in risk created by tamoxifen [13,14]. Because these two factors did not influence order effects, we will not discuss them at length here. However, as described below, we adjust for all five experimental factors in all our multivariate analyses.

The last debiasing factor we explored was the presence or absence of contextual information about competing health risks. In the *Guide to Decide*, all women learned their 5-year risk of breast cancer, with their average risk being 2.5%. Such information may be difficult to factor into decision making, given how unusual it is for people to learn about their 5-year risk of specific diseases. The way women incorporate this risk of breast cancer into their judgments, therefore, may depend on whether they are given additional information that helps them process this risk [15]. Contextual information might reduce order effects by giving people a standard of comparison by which to better process information about the risks and benefits of tamoxifen. With this standard of comparison, risk information might become more salient or more memorable, and thereby reduce people's susceptibility to order effects. In this study, we provided half of our participants, selected at random, with information on competing risks that they faced over the next 5 years—risks of experiencing colon cancer, a heart attack, or all-cause mortality. The remaining half, selected at random, did not receive any such contextual information.

By independently varying these factors – the order of risk and benefit information, and the presence or absence of contextual information about risk – we are able to test whether contextual information reduces order effects. The order with which people received information in our DA is illustrated in Fig. 1. As this figure shows, we presented contextual risk information to participants directly after giving them information on their 5-year risk of breast cancer. Thus, participants received contextual risk information before they had received any information about the risks and benefits of tamoxifen.

### 1.3. Outcome measures

#### 1.3.1. Knowledge

Knowledge was assessed with six questions about the risks and benefits of tamoxifen. Four questions focused on the risks

Fig. 1. Order of information participants received across experimental groups.

(increased incidence of endometrial cancer, hormonal symptoms, blood clotting problems, and cataracts) and two focused on the benefits (reduced chance of breast cancer and broken bones). Participants indicated who was more likely to experience each risk and benefit: (1) women who take tamoxifen, (2) women who do not take tamoxifen, (3) both groups are equally likely, and (4) do not know. Responses were scored as correct or incorrect (do not know was coded as incorrect) and correct responses were summed. The reliability coefficient (alpha) was .85.

1.3.2. Perceptions of the risks and benefits

After women completed the *Guide to Decide*, we asked them a series of questions exploring their subjective perceptions of the risks and benefits of tamoxifen. Wording of these items is shown in Table 2.

1.3.3. Analyses

To examine our hypotheses, we tested multivariate linear regression models using a two-step entry procedure. On the first step, we entered dummy variables representing main effects for each of the five experimental factors. This allowed us to examine whether the order manipulation influenced our dependent variables – knowledge or risk perceptions – independently of the other experimental manipulations. On the second step, we entered three interaction terms for the three potential debiasing factors: context, pictograph, and incremental risk. This procedure allowed us to examine whether order effects were reduced by any of these three factors. To aid in interpretation, where we observed statistically significant interaction terms, we present mean levels of our dependent variables broken down by each independent variable.

2. Results

2.1. Participants

Of the 8896 women who received an invitation letter, 1218 (14%) visited the website. Of these, 749 (61%) were eligible and 663

(89% of eligible) consented to participate. Ultimately, 632 participants (84%) completed the post-test. The DA and the post-test took an average of 49 min to complete.

Table 1 describes the sample’s demographic characteristics. Participants were on average 58.9 years old, White, and well-educated. Gail scores ranged from 1.7 to 17.3 ( $M = 2.56, SD = 1.26$ ).

Fig. 2 displays participants’ knowledge scores (percent of correct answers provided) as a function of whether they received risk information first or last, and whether they received risk context information. As shown on the left hand of the figure, there was a significant order effect for knowledge among those participants who did not receive risk context information; in this group, those who received risk information last had higher knowledge scores than those who received it first ( $p < .001$ ), likely because four of the six knowledge questions dealt with tamoxifen’s risks and this risk information was fresher in their minds. No such order effect was seen among participants who received risk context information ( $p = .53$ ).

Table 1  
 Demographic characteristics of sample.

Age
Mean: 58.86 (SD = 7.6)
Range = 40–74 <sup>a</sup>
Race
White: 94.9%
Black: 2.3%
Asian: 1.7%
Education
High school diploma or GED: 8.0%
Some college: 26.3%
Bachelor’s degree or higher: 65.7%
Gail score
Mean: 2.56 (SD = 1.26)
Range: 1.7–17.3

<sup>a</sup> Age range was constrained by eligibility requirements of study.

Fig. 2. Order effects in knowledge as a function of whether participant received information on risk context.

A linear regression confirmed the pattern suggested by the figure. In step one of this analysis, the dependent variable (percent correct answers) was significantly associated with the order of benefit and risk information, with higher knowledge among those who received risk information last ( $b = .07$  (.02),  $p = .003$ ).

In step two, we added the three interaction terms and found a significant interaction between order and risk context ( $b = -.10$  (.05),  $p = .04$ ), confirming that the order effect in knowledge was

significantly smaller for those participants who received information on risk contexts. The other two interactions (order by pictograph, and order by presentation of incremental risk information) were not significant ( $p = .55$ ,  $p = .68$ , respectively), indicating that neither of these factors reduced the biasing effect of order on knowledge.

Table 2 displays people's risk perceptions as a function of whether they received risk information first or last, and whether they received risk context information. As shown in the left hand portion of the table, there were several order effects in risk perceptions among those participants who did not receive risk context information; those participants who received risk information last were generally more worried about the side effects of tamoxifen and less impressed with tamoxifen's ability to prevent breast cancer. No such order effect was seen among participants who received risk context information.

Linear regressions confirm the pattern suggested by the table. As in the previous regression model, in step one of this analysis, for each of the measures of risk perceptions, we tested whether the dependent variable (risk perception) was significantly predicted by the order of benefit and risk information. We observed statistically significant effects of order for worry about side effects ( $b = .22$  (.08),  $p = .01$ ) and belief that side effects are common ( $b = .16$  (.07),  $p = .02$ ). In step two, we added interaction terms and found significant interactions between order and risk context information for both of these variables (for worry,  $b = -.39$  (.17),  $p = .02$ ; for belief that side effects are common,  $b = -.31$  (.14),  $p = .02$ ), confirming that the order effect in risk perceptions was significantly smaller among those people who had received information on risk context.

Table 2  
Order effects in risk perceptions as a function of whether participant received information on risk context.

Risk perceptions (1–5)	No risk context information			Risk context information		
	Risks last	Risks first	<i>p</i>	Risks last	Risks first	<i>p</i>
Taken all together, how worried would you be about getting any of the above health conditions if you did take tamoxifen? 1 = Not at all worried 5 = Extremely worried	3.6	3.1	.00	3.3	3.3	.82
Taken all together, how common do you think the above health conditions are for women who take tamoxifen? 1 = Not at all common 5 = Extremely common	3.0	2.7	.01	2.8	2.8	.96
If you were to choose to take tamoxifen, how likely do you think you would be to experience a side effect? 1 = Not at all likely 5 = Extremely likely	3.3	3.2	.15	3.3	3.3	.78
Compared to the average woman (your age and in your health), what do you think your chances are of developing breast cancer in the next 5 years? 0 = Much less than the average woman 3 = Same as the average woman 6 = Much higher than the average woman	2.3	2.4	.36	2.4	2.4	.79
If you were to choose to take tamoxifen, how likely do you think you would be to get breast cancer in the next 5 years? 0 = Not likely at all 10 = Extremely likely	1.7	2.1	.03	2.1	2.1	.94

### 3. Discussion and conclusion

#### 3.1. Discussion

Our study demonstrates, first and foremost, that the order in which people learn about the risks and benefits of healthcare interventions can influence their knowledge and perceptions of these same risks and benefits. Second, our study illustrates a method investigators can use to test for biases such as order effects—by randomizing participants to receive risks and benefits in varying orders. Third, our study illustrates a method to test for factors that could potentially eliminate or reduce such biases, by conducting an experiment testing potential debiasing factors.

DAs deserve an increasingly common role in healthcare practice. Patients frequently face preference-sensitive decisions without enough information about the pros and cons of their healthcare alternatives, and without a formal recognition that their preferences are an important part of the decision. Our *Guide to Decide* provided information about tamoxifen to women who faced a high risk of being diagnosed with first breast cancer in the next 5 years. (Our revised *Guide to Decide* now also informs women about raloxifene, based on the results of the STAR trial.) Like many other well-designed DAs, the *Guide to Decide* succeeds in communicating the risks and benefits of Tamoxifen to women in a manner that helps them make an informed decision.

But knowledge, while necessary for good decision making, is not always sufficient [16]. As this study demonstrates, the same information presented in a different order can lead to different perceptions. This kind of order effect threatens the goals of DA developers, to inform patients about healthcare alternatives in a neutral manner. Given such order effects, DA developers need to worry that their choices about how to present the information – for example risks before benefits or vice versa – can unduly influence people’s perceptions of their healthcare alternatives.

Some readers may wonder what we mean when we refer to order effects as a “cognitive bias.” We use this term in the sense it has been used in the judgment and decision making literature, as best represented in the seminal work of Kahneman and Tversky who, in the 1970s, documented a host of heuristics (cognitive shortcuts) that lead people to make flawed judgments [17,18]. According to this tradition, a judgment is biased if it departs from accepted norms of rationality [19]. In the case of order effects, because the risk/benefit ratio of Tamoxifen is the same regardless of whether a person receives risk or benefit information first, so too should be their knowledge of these risks and benefits, as well as their perceptions of tamoxifen’s risk/benefit ratio. Any departure from this equivalence norm is characterized, therefore, as a bias.

But is such a bias necessarily bad? Some experts may question whether order effects are a threat to DA developers, wondering instead whether the psychological influence of recency can be used to improve DA design. For instance, according to this reasoning, DA developers could place risk information near the end of their materials, to make sure people remember such information at the time of their decision. But this approach only succeeds if DA developers are convinced that risk information is, somehow, more relevant to a person’s decision than other informations, such as information about the benefits of a specific treatment alternative. But making such a determination, about the priority of risk information over other informations, will no doubt strike some experts as a value judgment that departs from the neutrality that DA developers typically strive for in developing their materials.

Clearly, part of the challenge in designing any DA is to determine which information is most important for decision makers to understand. The order effect we identified in this study suggests that it can be difficult to find a way of presenting such information to decision makers that enables them to place equal

weight on equally important information. Fortunately, our experiment also suggests a way to avoid this problem—by implementing a DA design that reduces order effects.

Is risk context information the solution to this problem? We think such a conclusion would be premature. The order effects seen in this study are relatively small and the elimination of such order effects by risk context information needs to be replicated in other studies before such an approach could be deemed a success. We do not have an adequate understanding, on the basis of the current study, for why contextual information about competing risks eliminated the order effect we discovered in this study. In future research, it will be important to determine how common such order effects are, and then to develop a better understanding for how to reduce or eliminate such order effects. Our goal in this article is not to convince DA developers that we have solved this specific problem. Instead, our goal is to encourage researchers to focus more efforts on exploring biases such as order effects, as well as testing methods to eliminate such biases.

Our study has several limitations that warrant caution about how to interpret our results. First, our sample was comprised mostly of white women, recruited from only two locations in the United States. In addition, like many other Internet studies, ours was plagued by a low response rate. Thus the generalizability of our results is a concern. Nevertheless, the lack of generalizability of our study does not negate our experimental finding. Because we randomized participants across the versions of our DA, our order effect and our debiasing of the order effect were shown to exist for this population.

Second, the order effects we found were small, and not necessarily of clinical significance. Nevertheless, it is possible that order effects will be larger in other contexts. In addition, more important than the size of our order effects, was our desire to show how randomized experiments can help uncover decision lowoun-13.16.7ct priorityo9ion,hc(co)90ct prioriscussi(on102(c

### 3.3. Practice implications

Clinicians need to be aware that DAs which they use in clinical setting may introduce cognitive biases. In the future, they can hopefully be able to rely on DAs that have been thoroughly tested for such biases.

### Acknowledgments

We would like to thank Drs. Victor Strecher, Priti Shah, and Mick Couper for their assistance in the design of the study. The Center for Health Communication Research (especially Michael Nowak) did an excellent job turning our decision aid into a well-designed and easily navigated web site. We would also like to thank Roy Pardee and Rick Krajenta for their assistance with compiling the databases of eligible participants, and Dan Hayes and Paula Lantz for help designing the decision aid. We express our gratitude to the Department of Radiology at Henry Ford Health System, particularly Dr. Matthew Burke, for their help in identifying eligible patients. Finally, we are very grateful to all the women who participated in this study.

### References

- [1] O'Connor AM, Rostom A, Fiset V, Tetroe J, Entwistle V, Llewellyn-Thomas H, Holmes-Rovner M, Barry M, Jones J. Decision aids for patients facing health treatment or screening decisions: systematic review. *Brit Med J* 1999;319:731–4.
- [2] O'Connor AM, Fiset V, DeGrasse C, Graham ID, Evans W, Stacey D, Laupacis A, Tugwell P. Decision aids for patients considering options affecting cancer outcomes: evidence of efficacy and policy implications. *J Natl Cancer Inst Monogr* 1999;25:67–80.
- [3] Ubel P. Beyond knowledge: figuring out how to help people make “good” decisions. In: Shafir E, editor. *The behavioral foundations of policy*: princeton university and russell sage foundation presses; in press.
- [4] Redelmeier DA, Rozin P, Kahneman D. Understanding patients' decisions. Cognitive and emotional perspectives. *J Am Med Assoc* 1993;270:72–6.
- [5] Ubel PA. Is information always a good thing? Helping patients make “good” decisions. *Med Care* 2002;40:V39–44.
- [6] Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst* 1998;90:1371–88.
- [7] Trotman K, Wright A. Order effects and recency: where do we go from here? *Account Financ* 2002;40:169–82.
- [8] Molenaar S, Sprangers MA, Postma-Schuit FC, Rutgers EJ, Noorlander J, Hendricks J, de Haes HC. Feasibility and effects of decision aids. *Med Decis Making* 2000;20:112–27.
- [9] Nelson WL, Han PK, Fagerlin A, Stefanek M, Ubel PA. Rethinking the objectives of decision aids: a call for conceptual clarity. *Med Decis Making* 2007;27:609–18.
- [10] Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989;81:1879–86.
- [11] Nair N, Strecher V, Fagerlin A, Resnicow K, Murphy SA, Little R, Chakraborty B, Zhang A. Screening experiments and the use of fractional factorial designs in behavioral intervention research. *Amer J Public Health* 2008;98:1354–9.
- [12] Denes-Raj V, Epstein S, Cole J. The generality of the ratio-bias phenomenon. *Pers Soc Psychol Bull* 1995;21:1083–92.
- [13] Zikmund-Fisher BJ, Ubel PA, Smith DM, Derry HA, McClure JB, Stark A, Pitsch RK, Fagerlin A. Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. *Patient Educ Couns* 2008;73:209–14.
- [14] Zikmund-Fisher B, Fagerlin A, Roberts T, Derry H, Ubel P. Alternate methods of framing information about medication side effects: incremental risk versus total risk occurrence. *J Health Commun* 2008;13:107–24.
- [15] Windschitl PD. Judging the accuracy of a likelihood judgment: the case of smoking risk. *J Behav Decis Making* 2002;15:19–35.
- [16] Ubel P. Free market madness: why human nature is at odds with economics—and why it matters. Boston: Harvard Business Press; 2009.
- [17] Tversky A, Kahneman D. Availability: a heuristic for judging frequency and probability. *Cognitive Psychol* 1973;5:207–32.
- [18] Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974;185:1124–31.
- [19] Baron J. Thinking and deciding. 2nd ed., New York: Cambridge University Press; 1994.