This is supplementary material to an Accepted Manuscript of an article published by Taylor & Francis in *Psychology & Health* on [accepted July 2018—IN PRESS], that will be available online: http://www.tandfonline.com/10.1080/08870446.2018.1508685.

Article citation:

MacFarlane, D., Hurlstone, M., & Ecker, U. (in press) Reducing Demand for Ineffective Health Remedies: Overcoming the Illusion of Causality. *Psychology & Health*. [Supplemental Material] doi:10.1080/08870446.2018.1508685

Supplementary Material for 'Reducing Demand for Ineffective Health Remedies: Overcoming the Illusion of Causality'

Douglas MacFarlane[a]*, Mark J. Hurlstone[b], and Ullrich K. H. Ecker[c]

[abc]*School of Psychological Science, University of Western Australia, Perth, Australia;*

*Corresponding author: School of Psychological Science, University of Western Australia; M304, 35 Stirling Hwy, Perth 6009, Australia; phone: +618 6488 1418; email: douglas.macfarlane@research.uwa.edu.au; ORCiD: 0000-0003-4378-9294

[b] School of Psychological Science, University of Western Australia; M304, 35 Stirling Hwy, Perth 6009, Australia; phone: +618 6488 3249; email: mark.hurlstone@uwa.edu.au;

[c] School of Psychological Science, University of Western Australia; M304, 35 Stirling Hwy, Perth 6009, Australia; phone: +618 6488 3257; email: Ullrich.ecker@uwa.edu.au; ORCiD: 0000-0003-4743-313X

Supplementary Material for 'Reducing Demand for Ineffective Health Remedies: Overcoming the Illusion of Causality'

***Objective*:** This supplementary document reports the results of a pilot study conducted prior to the main study. The primary aim of the pilot study was to assess the convergent validity of the auction mechanism and four predictors of participants' willingness to pay for multivitamins: estimated usage, efficacy belief, general-attitudes, and irrational health beliefs. A secondary aim was to assess the internal consistency and structure of an attitudinal scale designed to measure participants' pre-existing general attitudes to health supplements. Another secondary aim was to establish a baseline with regards to what factors people perceive to be important when making decisions about health purchases.

***Design*:** The pilot study was conducted online with US participants ($N = 194$) recruited via Amazon's Mechanical Turk. The same measures and procedure used in the main study were used in the pilot study, with a few exceptions. First, participants were not subject to any interventions in the pilot study. Second, the pilot study measured people's *hypothetical* willingness to pay (WTP) for the effervescent multivitamin product using the same auction mechanism employed for the first auction in the main study.

Results: We obtained convergent evidence that the auction mechanism indexes consumer demand for multivitamin supplements via positive correlations between hypothetical WTP and estimated usage, efficacy belief, and general attitudes. Second, we found strong evidence that individuals are vulnerable to the illusion of causality and place relatively little value on clinical information when making health purchases.

Conclusion: The pilot study confirmed that our survey instruments were suitable for answering the research question addressed in the main study.

Keywords: causality; consumer behaviour; intervention; health communication; health behaviour; health education;

Pilot study

Prior to conducting the main study, a pilot study was conducted online with US participants recruited via Amazon's Mechanical Turk. The study was launched and completed on the 23rd of February 2017. Participants in this study were not subjected to any interventions. Instead, the primary aim was to assess the convergent validity of the auction mechanism and four predictors of participants' hypothetical willingness to pay for multivitamins: estimated usage, efficacy belief, general-attitudes, and irrational health beliefs. A secondary aim was to assess the internal consistency and structure of the general attitudes scale developed to measure participants' pre-existing general attitudes to health supplements. Another secondary aim of the pilot study was to establish a baseline with regards to what factors people perceive to be important when making decisions about health purchases.

The pilot study was the same as the main study, with the following exceptions. First, whereas the main study used two incentivized auctions in which participants could win real health products by bidding with real money, the pilot study used a hypothetical version of the first auction reported in the main study. Second, the pilot study included an additional measurement scale, namely the Irrational Health Belief Scale (Christensen, Moran, & Wiebe, 1999), which purportedly measures the degree to which individuals hold irrational beliefs about health products. Third, the main study included an additional predictor, namely participants' current health (participants were asked to indicate their current health on a 4-point scale; 1 = fine, healthy; 4 = sick, exhausted) that was not used in the pilot study.

Design

Ethics approval to conduct the study was granted by the Human Ethics Office of the University of Western Australia in accordance with the requirements of the Australian

National Statement on Ethical Conduct in Human Research (NHRMC, 2007).

Participants

In total, 201 participants responded to the survey. Several a-priori exclusion criteria were used (details are provided in the results section). Based on these criteria, seven participants were removed for careless responding. The final sample thus included N = 194 participants (81 females, 113 males; age M = 36.71 years; SD = 10.97). The number of participants was pre-determined based on the recommended numbers for score validation methods such as a principal components analysis (Field, Miles, & Field, 2012). Pay was determined at a rate of USD$7.40/hour pro rata, above US federal minimum wage, which has been suggested as appropriate to attract 'normal' workers (Rouse, 2015).

A pre-exclusion criteria was also set prior to recruitment that required participants to have (i) a completion number of at least 5,000 'Human Intelligence Tasks' on Mechanical Turk, and (ii) an approval rate of greater than 97%. These qualifications are recommended to ensure quality responses from Mechanical Turk workers (Peer, Vosgerau, & Acquisti, 2014).

Materials

Predictors 1 and 2: Multivitamin consumption

Participants were asked whether they had taken multivitamins in the past, and if so, to then provide an estimated previous usage frequency on an eight-point scale (1 = *not in the past few years*, 8 = *every day*). Participants then rated their efficacy belief, that is, their belief in the effectiveness of the routine consumption of multivitamins for maintaining general health, on a 5-point scale (1 = *not effective at all*, 5 = *extremely effective*; a sixth point was included in the scale so that participants could respond 'I don't know').

Predictor 3: General attitudes toward health supplements and alternative medicines

General attitudes toward health supplements and alternative medicines were assessed using an 18-item scale (see Table 1). Each item consisted of a statement that was based on a literature review of the motivations shown to influence the consumption of alternative health products (e.g., 'Vitamins are natural and supplements are therefore safe'). Participants were asked to rate how much they agreed with each statement using a 5-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*). A composite score was calculated for each participant that indicated their general attitudes to health supplements and alternative medicines (hereafter referred to as the 'general-attitude score'). To measure response consistency, each item was paired with a reverse-phrased statement of similar meaning (i.e., 9 pairs of items). The order of items in the general-attitude survey was randomized to control for order effects.

Predictor 4: Irrational health belief scale

The irrational health belief scale is a 20-item survey designed to assess individual differences in the tendency to engage in health-related cognitive distortions (Christensen et al., 1999). Each survey item describes a health situation and a corresponding thought response (e.g. 'You have been taking a medication for six months and your medical problem has not improved. Your doctor has suggested a new drug. You think to yourself, "If the last medication didn't help, a new one won't do any good."'). Participants are asked to imagine that the situation applies to them and then indicate how similar the thought is to their own thought pattern in that situation using a 5-point scale (1 = *not at all like I would think,* 5 = *almost exactly like I would think).* The order of items in the irrational health belief scale was randomized to control for order effects. Collecting responses to both the irrational health belief scale and the general-attitude survey provided means for assessing each scale's construct validity.

Willingness-to-pay (WTP)

To assess WTP, data was collected using a hypothetical version of the Becker-Degroot-Marschak (BDM) auction mechanism (Becker, Degroot, & Marschak, 1964; Thrasher, Rousu, Hammond, Navarro, & Corrigan, 2011). Participants were asked to imagine they had been given $5 with which to place a bid on a tube of effervescent multivitamin tablets. They were shown a plain-packaged picture of the product and provided with some descriptive text about multivitamins, including some common health claims and a popular pseudo-scientific causal explanation as to why supplements are thought to provide health benefits. The exact text was:

> 'Vitamins and minerals are micro-nutrients. They are, in small amounts, essential for health. The lack of a specific micro-nutrient may cause or predispose someone to disease. Micro-nutrient supplements are widely available; they are usually referred to as "multivitamins." The health benefits of multivitamins have been claimed to include a reduction in cardiovascular disease and cancer, as well as an improvement in cognitive function. Such benefits are commonly explained by stating that multivitamin supplements boost the body's natural immune system.'

Participants were asked to bid only the amount that would reflect how much they would be willing to pay for that product. Participants were told that this was different to other auctions in that they could only bid once, and that it was in their best interest to only bid the amount they were willing to pay. Participants were required to enter their bid amount b in cents $b \in (0, 500)$. They knew that this amount would be compared against a random number $r \in (0, 500)$ drawn from a uniform distribution, and that if $b \geq r$ then they would win the auction and hypothetically purchase the product for amount b but keep $500 - b$ of their imagined endowment; otherwise if $b < r$ then they would lose the auction but hypothetically keep the full $5 endowment. The decision for participants to pay the bid price, as opposed to the random price, which is the standard BDM model,

was based on previous laboratory experience that participants find paying the random price confusing. We suggest the approach taken in the present paper is more intuitive as it reflects the more familiar and real-world format adopted by online auction websites, whereby consumers pay the bid amount once the reserve has been exceeded. Participants knew that the auction and endowment were hypothetical and that they would not receive any money, nor the multivitamin product on which they were bidding, if their bid was successful. Prior to the main auction, participants were given the chance to participate in two practice auctions using a hypothetical $1 endowment to bid on a bottle of water.

Fictitious efficacy rating

This measure introduced participants to a fictional nausea drug 'Product Z.' To create an illusion of causality, participants were only given a half-contingency table, which showed that 4 out of 5 people experienced a health benefit from taking Product Z. Participants were asked to rate the effectiveness of Product Z as either *not effective*, *mildly effective*, *very effective*, or to indicate there was *insufficient information* to make a determination of effectiveness. This question was designed to produce an estimate of the proportion of people that fall prey to the illusion of causality, as indexed by the proportion of individuals who failed to identify there was insufficient information to determine causality.

Factors influencing future health purchases

To establish a baseline of the factors that people perceived as important when making future health-related purchases, participants were asked to rate the importance of 15 health-related factors using a 5-point Likert scale (1= *not important at all*, 5 = *extremely important*). For the purposes of this pilot study, only three items were considered of interest. The most critical item was placebo comparison ('the number of people who

did, and did not, experience a benefit when taking a sugar pill'). Placebo comparison information is critical for making rational health purchases, because it is necessary for assessing the efficacy of a remedy compared to taking a placebo. The other two items of interest were product comparison ('the number of people who did, and did not, experience a benefit after taking a product'), and 'known side effects.' We reasoned that both product comparison and known side effects would be important to people, and thus would provide a suitable baseline from which to compare participants ratings of placebo comparison information. The remaining 12 items were distractor items (e.g., importance of 'advertising claims'); item order was randomized.

Procedure

Participants were recruited via a survey link provided through Mechanical Turk. At the start of the survey, participants were provided with an information form and provided informed consent. They then responded to questions on demographics, multivitamin consumption, and the general-attitude scale. Next, participants were shown the multivitamin product and asked to indicate their WTP for it via the hypothetical BDM auction. Participants then responded to the fictional product question and the questions regarding their future health purchases. Finally, they completed the irrational health belief scale.

Objectives

We made three predictions. *Hypothesis* I. Higher estimated usage, efficacy belief, and general-attitude scores would be associated with higher WTP for multivitamins. *Hypothesis* II. Higher irrational health belief scale scores would be associated with higher WTP for multivitamins. *Hypothesis* III. Most participants would not recognize the value of clinical trial information when making future health purchases. Specifically, most people would (1) fail to recognize that there was insufficient

information to assess the efficacy of the fictional product, and (2) the placebo-comparison factor would be rated as less important than the product-comparison or side-effects factors.

Data analysis

Most analyses were conducted in R (R Core Team, version 3.3.2) with the exception of the Principal Components Analysis, which was conducted in SPSS (IBM Corp, version 22). To assess the internal structure of the general-attitude survey, a principal components analysis was conducted on the 18-items with oblique rotation (direct oblimin). Sampling adequacy for the analysis was verified by the Kaiser-Meyer-Olkin (KMO) measure. Bartlett's test of sphericity was used to assess whether correlations between items were sufficiently large for a principal components analysis. An initial analysis was run to determine how many components to extract with the principal components analysis. To establish internal consistency reliability, we computed Cronbach's alpha.

To assess whether estimated usage predicted WTP, we excluded participants who indicated they had never taken multivitamins before (since they were not asked to estimate their usage). To assess whether efficacy belief predicted WTP, we excluded participants who answered, 'I do not know.' Associations between WTP and each of the four potential predictors: estimated usage, efficacy belief, general-attitude score, and irrational health belief scale were analysed using Spearman's correlation coefficient (Hypotheses 1 and 2). To test whether participants showed an overall preference for certain responses on the effectiveness of the fictional product, we performed a multinomial goodness-of-fit test to assess whether the response counts differed from chance. Planned comparisons were conducted using binomial tests with sequential Bonferroni adjustment. To test whether participants' ratings of the importance of the

three factors for future health purchases (product comparison, placebo comparison, and side effects) were significantly different, a Friedman's ANOVA was conducted. Planned comparisons were then conducted using Wilcoxon sign-ranked tests with sequential Bonferroni adjustment.

Results

Exclusion criteria

A-priori criteria were applied to identify careless responders based on recommendations from the literature (Lowry, D'Arcy, Hammer, & Moody, 2016; Oppenheimer, Meyvis, & Davidenko, 2009). Participants were excluded who (i) gave non-differentiated answers to every question in a survey block (Barge & Gehlbach, 2011; Hamby & Taylor, 2016); (ii) failed an attention-trap question (Lowry et al., 2016); (iii) completed a survey block in less than the allocated minimum reading time (i.e., > 600 words per minute; Carver, 1985); and (iv) whose responses were, on average, overly inconsistent between pairs of equivalent questions (i.e., an odd/even threshold of > 2 Likert points apart; Curran, 2016). The exclusion criteria, and subsequent data analysis, were applied separately to the three sections of the experiment: (i) the general-attitude scale (n = 7), (ii) questions about future health purchases (n = 9), and (iii) the irrational health belief scale (n = 42). Exclusion criteria were applied separately because each section varied in cognitive demand, which is known to increase careless responding (Krosnick, 1991). Further, as the irrational health belief scale was the longest and most demanding section, it was deliberately placed at the end of the study to reduce careless responding in the other sections. The same exclusion criteria were adopted in our main study with one exception. Specifically, we removed the attention-trap questions from our a-priori exclusion criteria following research suggesting that such items can increase socially desirable responding (Clifford & Jerit, 2015).

Internal consistency and structure of general-attitude scale

Reverse-phrased items were reverse coded for analysis, so that higher total general-attitude scores indicated more favourable attitudes toward health supplements. For the principal components analysis, sampling adequacy was verified, KMO = .885 (Hutcheson & & Sofroniou, 1999), and all KMO values for individual items were > .8, which is above the acceptable limit of .5 (Field et al., 2012). Correlations were found to be sufficiently large for a principal components analysis, χ^2 (153) = 1766.68, $p < .001$. Five components had eigenvalues over Kaiser's criterion of 1 and in combination explained 68.4% of the variance. The scree plot showed two points of inflexion, after components one and five, suggesting that either one or five components may be appropriate (Field et al., 2012) . Table 1 shows the component loadings after rotation. The items that cluster on the same components suggest that component 1 represents a general health factor, component 2 relates to safety, component 3 to supplements, component 4 to alternative medicine, and component 5 to nutrition. The general-attitude survey was found to be associated with a high level of internal consistency, Cronbach's $\alpha = .91$.

Predictors of willingness to pay

Figures 1, 2, and 3 show WTP as predicted by the three predictor measures of estimated usage of multivitamins, efficacy belief, and general-attitude score, respectively. Correlations between the predictors of interest and WTP provided statistical confirmation of Hypothesis I. Specifically, WTP was positively associated with estimated usage of multivitamins, $N = 170$, $r_s = .25$, $p < .001$, efficacy belief, $N = 185$, $r_s = .31$, $p < .001$, and general-attitude scores, $N = 194$, $r_s = .42$, $p < .001$. However, we

did not find support for Hypothesis II as irrational health belief scale scores were not significantly associated with WTP, $N = 159$, $r_s = .10$, $p = .21$.[1]

Future health purchases

Figure 4 shows participants' responses to the fictional product question. These responses were found to be significantly different to the pattern expected by chance, $N = 194$, $p < .001$. Post-hoc tests revealed that all four response frequencies were significantly different from the expected value of 48.5 ($N/4$): 'Very effective' was higher than the expected value $p < .001$; 'Insufficient evidence' was lower, $p < .001$; 'Not effective' was lower, $p < .001$; and 'Mildly effective' was lower, $p < .05$. The results showed that the majority of people succumbed to the illusion of causality, since the majority (69%) responded that the product was very effective and only a minority (12 %) correctly identified that there was insufficient information to determine efficacy.

Figure 5 shows results of participants' ratings of factors that may influence future health purchases. The results provided statistical confirmation of Hypothesis III. Specifically, we found that participants' ratings of the importance of the three items of interest (product comparison, placebo comparison, and side effects) to be significantly different, $N = 194$, $\chi^2 (2) = 116.3$, $p < .001$. As expected, placebo comparison ($Mdn = 3$, 'Moderately Important') was rated significantly less important than product comparison ($Mdn = 4$, 'Very Important'), $p < .001$, $r = -.43$, or side effects ($Mdn = 5$, 'Extremely Important'), $p < .001$, $r = -.65$. Finally, product comparison was rated significantly less important than side effects, $p < .001$, $r = -.46$.

[1] All analyses were also conducted with careless responders retained. These results showed that the exclusion criteria made a negligible difference to the results.

Discussion

The pilot study employed a variation of the BDM auction mechanism in a hypothetical scenario to examine the relationship between people's WTP for multivitamins and their (i) estimated usage habits, (ii) belief in the efficacy of multivitamins, and (iii) general attitudes towards dietary supplements and alternative medicines. We found that more frequent estimated usage, greater efficacy belief, and more favourable general attitudes reliably predicted a greater WTP for multivitamin products (Hypothesis I). The pilot study also examined the relationship between people's irrational health belief scale scores and their WTP for multivitamins. Contrary to our prior prediction, we found that the irrational health belief scale was not a reliable predictor of people's WTP (Hypothesis II). The pilot study also sought to establish a baseline for people's propensity to recognize the value of clinical trial information when making future health purchases. The results showed that when provided with only half a contingency table for a fictitious health product, most people succumbed to the illusion of causality by incorrectly responding that the information presented indicated that the product was 'very effective.' The results also showed that placebo-comparison information was rated as relatively unimportant, despite it being critical to determine if a product is efficacious (Hypothesis III).

Validity of the general-attitude scale and the auction mechanism

The primary aim of the pilot study was to demonstrate the validity of both the general-attitude survey and the auction mechanism. Two aspects of the data suggest that the general-attitude scale is a valid measure of pre-existing attitudes toward supplements. Specifically, the general-attitude scale was found to be highly internally consistent and it was shown to measure five distinct but related components of attitudes towards supplements. This result was expected because the general-attitude scale was based on

previous research into the underlying drivers of the consumption of alternative medicines and dietary supplements. For example, items were based on findings from a study (Barnes, Ball, Desbrow, Alsharairi, & Ahmed, 2016) indicating that the two most common reasons for taking multivitamins were to supplement nutrition (60%) and to increase immune strength (49%). The finding that all three moderators predicted WTP suggests that this variation of the BDM auction mechanism, is a valid tool for measuring consumer demand for multivitamin products. These findings validate our methodology as suitable for answering the central question posed in the main study, namely whether or not the intervention treatment leads to a measurable behaviour change.

Validity of the irrational health belief scale

A secondary goal of the pilot study was to test the validity of the irrational health belief scale. The finding that the irrational health belief scale did not predict WTP raises questions about the validity of this scale as a measure of irrational health beliefs. Taken in isolation, this conclusion is questionable. For example, one might alternatively conclude that it is the validity of the auction mechanism itself, as a measure of consumer demand for multivitamins, that should be called into question. However, the fact that scores on the general–attitude scale—which also features items designed to tap irrational health beliefs—did reliably predict WTP in the auction confers support for the first conclusion. For these reasons, we decided to omit the irrational health belief scale from the main study.

One possible reason for the failure of the irrational health belief scale to predict WTP is that people who ascribe to certain irrational health beliefs may not be acting irrationally, instead they may be acting rationally with imperfect information. In other words, for any one individual, the available evidence concerning the efficacy of a health

belief—including personal experience, exposure to marketing claims, and anecdotes from one's own social network—may all *logically* lead to the conclusion that a health behaviour caused an unrelated beneficial outcome. Thus, there may be no irrational cognitive distortions driving an irrational health behaviour other than nescience. The same conclusion could not be drawn for the general-attitude scale since this measure included several items that directly measured peoples' attitude towards dietary supplements (e.g. 'Taking vitamin supplements is a good way to maintain general health').

Baseline of perceived importance of factors for making health purchases

The pilot study also established two baselines for comparing whether the effect of our interventions in the main study would generalize to future health purchases. The first baseline is that most people succumb to the illusion of causality, as indicated by responses to the fictional product. This supports previous research demonstrating the strength of this illusion (Matute et al., 2015; Yarritu & Matute, 2015). The second baseline is that placebo-comparison information was rated as less important than product-comparison information, which in turn was rated as less important than information about side effects. One explanation for these preferences is that people's perception of the importance of information is negatively correlated to its complexity, especially when making fast decisions. In other words, because placebo comparison information is both complex to obtain and complex to understand, this information is perceived as less important for everyday decisions. This supports previous research that many decisions are made using heuristics that 'ignore part of the information, with the goal of making decisions more quickly, frugally, and more accurately than more complex methods' (Gigerenzer & Gaissmaier, 2011). While simple heuristics can be efficient and sufficiently accurate for many types of decisions (e.g., Dhami, 2003;

Dhami & Harries, 2001), they can also lead to significant errors, especially when complex information is critical for inferring causality, as is often the case when the outcomes of a cause are not immediately known. This is particularly true of many health behaviours, as the benefits and/or side effects may not surface for many years and can differ considerably between individuals. These results further highlight the need for interventions that assist people in overcoming nescience.

Concluding remarks

There were two key findings of this pilot study. First, we obtained convergent evidence that the auction mechanism indexes consumer demand for vitamin supplements in the form of positive correlations between hypothetical WTP and the predictor measures of estimated usage, efficacy belief, and general attitudes. Second, based on the fictional health product scenario and future health purchases questionnaire, we obtained strong evidence that individuals are vulnerable to the illusion of causality, and that they place relatively little emphasis on clinical information when making health purchases. Thus, the pilot study confirmed that the survey instruments are suitable for answering the research question addressed in the main study.

References

Barge, S., & Gehlbach, H. (2011). Using the theory of satisficing to evaluate the quality of survey data. *Research in Higher Education, 53*(2), 182-200. doi:10.1007/s11162-011-9251-2

Barnes, K., Ball, L., Desbrow, B., Alsharairi, N., & Ahmed, F. (2016). Consumption and reasons for use of dietary supplements in an Australian university population. *Nutrition, 32*, 524-530. doi:10.1016/j.nut.2015.10.022

Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9*(3), 226.

Carver, R. P. (1985). How good are some of the world's best readers? *Reading Research Quarterly, 20*, 389-419. doi:10.2307/747851

Christensen, A. J., Moran, P. J., & Wiebe, J. S. (1999). "Assessment of irrational health beliefs: Relation to health practices and medical regimen adherence": Correction. *Health Psychology, 18*, 220-220. doi:10.1037/0278-6133.18.3.220

Clifford, S., & Jerit, J. (2015). Do attempts to improve respondent attention increase social desirability bias? *Public Opinion Quarterly*, 79, 790-802. doi:10.1093/poq/nfv027

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology, 66*, 4-19. doi:10.1016/j.jesp.2015.07.006

Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science*, 14(2), 175-180.

Dhami, M. K., & Harries, C. (2001). Fast and frugal versus regression models of human judgement. *Thinking & Reasoning, 7*, 5-27. doi:10.1080/13546780042000019

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R* [Google Books]. Retrieved from https://play.google.com/store/books/details?id=Q9GCAgAAQBAJ&hl=en

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451-482. doi:10.1146/annurev-psych-120709-145346

Hamby, T., & Ickes, W. (2015). Do the readability and average item length of personality scales affect their reliability? *Journal of Individual Differences*, *36*, 54-63.

Hamby, T., & Taylor, W. (2016). Survey satisficing inflates reliability and validity measures: An experimental comparison of college and Amazon Mechanical

Turk samples. *Educational and Psychological Measurement, 76*, 912-932. doi:10.1177/0013164415627349

Hutcheson, Graeme & Sofroniou, Nick (1999). *The multivariate social scientist: introductory statistics using generalized linear models.* Sage Publications, London.

IBM SPSS Statistics for Windows (Version 22)[Computer software]. Armonk, NY, IBM Corp.

Krosnick, J. A. (1991). Response strategies for coping with cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213-236.

Lowry, P. B., D'Arcy, J., Hammer, B., & Moody, G. D. (2016). "Cargo Cult" science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels. *The Journal of Strategic Information Systems, 25*, 232-240. doi:10.1016/j.jsis.2016.06.002

Matute, H., Blanco, F., Yarritu, I., Díaz-Lago, M., Vadillo, M. A., & Barberia, I. (2015). Illusions of causality: How they bias our everyday thinking and how they could be reduced. *Frontiers in Psychology, 6*, 1-14.

National Health and Medical Research Council. (2007). *National statement on ethical conduct in human research* (2007) - Updated December 2013 (the National Statement), Commonwealth of Australia, Canberra.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*, 867-872. doi:10.1016/j.jesp.2009.03.009

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46*, 1023-1031. doi:10.3758/s13428-013-0434-y

R Core Team (Version 3.3.2) [Computer Software] R: A language and environment for statistical computing. Vienna, Austria. R Foundation for Statistical Computing.

Rouse, S. V. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior, 43*, 304-307. doi:10.1016/j.chb.2014.11.004

Thrasher, J. F., Rousu, M. C., Hammond, D., Navarro, A., & Corrigan, J. R. (2011). Estimating the impact of pictorial health warnings and "plain" cigarette packaging: Evidence from experimental auctions among adult smokers in the United States. *Health Policy, 102*, 41-48. doi:10.1016/j.healthpol.2011.06.003

Yarritu, I., & Matute, H. (2015). Previous knowledge can induce an illusion of causality through actively biasing behaviour. *Frontiers in Psychology, 6*, 389.

Table 1. Factor loadings for principal components analysis with oblimin rotation on the general attitudes towards health supplements scale.

Item	*General Health*	*Additives / Experience*	*Nutrition*	*Alternative Medicine*	*Safety*
A person who is generally healthy can still benefit from taking vitamin supplements	0.89				
Taking vitamin supplements is a good way to maintain general health	0.87				
Vitamin supplements are a good investment	0.70				
Vitamin supplements are not worth the money	0.69				
Vitamin supplements increase immune strength	0.64				
I don't believe supplements increase immune strength	0.60				
The population is healthiest when people have access to both conventional and alternative medicines	0.50				-0.45
Vitamins are natural and supplements are therefore safe	0.43				
Vitamin supplements are only useful if a person has a specific deficiency	0.40				
I do not purchase products simply because they contain added vitamins		0.90			
Sometimes I choose certain products because they contain added vitamins		0.76			
The best way to know if a dietary supplement or medicine is effective is to try it for yourself		0.41			
Vitamin supplements are required because a typical diet is insufficient for optimal health			-0.80		
Vitamin supplements are not necessary for optimal nutrition.			-0.71		
One does not need to take vitamin supplements to maintain general health			-0.70		
Supplements or medicines should only be taken if there is scientific evidence that they work				0.86	
Most alternative medicines are ineffective		0.46		0.50	
Taking vitamin supplements may pose significant health risks					0.79
Eigenvalues	7.63	1.28	1.25	1.11	1.04
% of total variance	42.39	7.11	6.94	6.17	5.78
α	0.90	0.74	0.80	0.47	0.16

Note. Factor loadings <.40 are not shown. Data excludes seven careless responders (N=194).

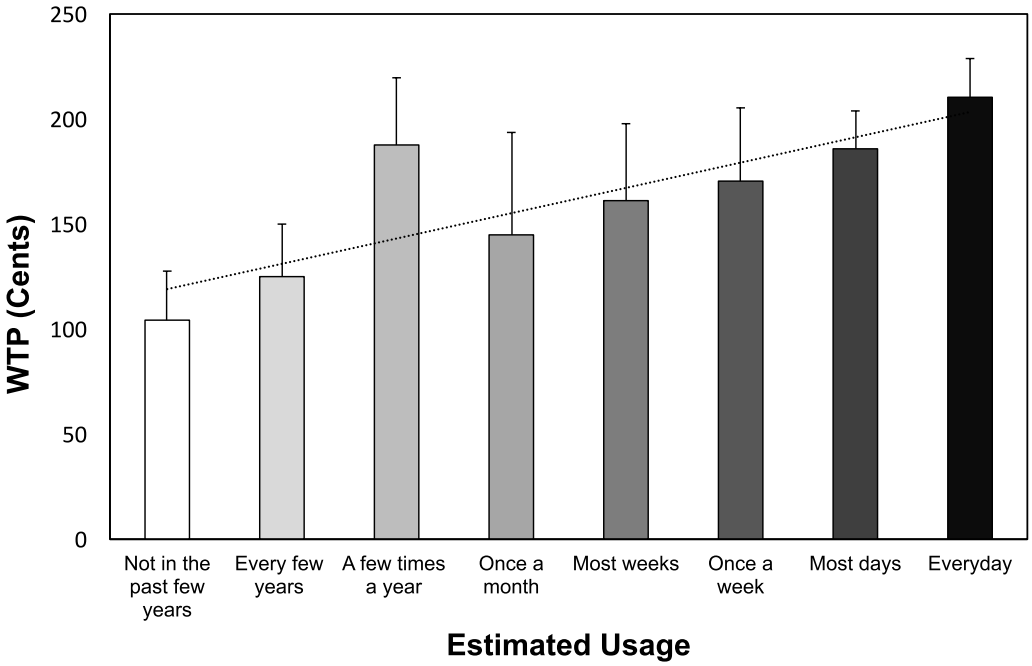

Figure 1.

Participants' estimated frequency of their usage of multivitamins against WTP for the multivitamin product, $N = 170$. Error bars indicate standard errors.

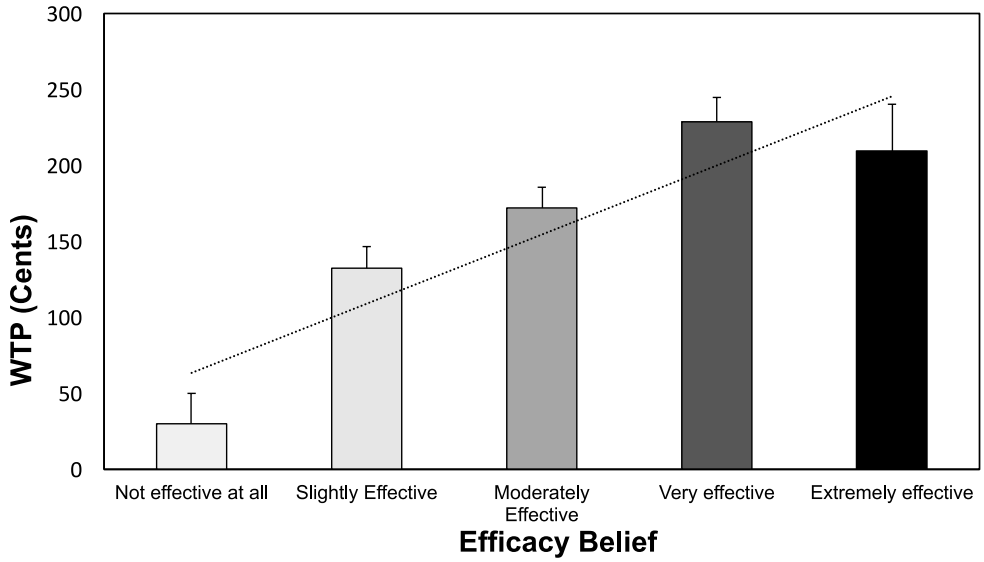

Figure 2.

Participants' estimated belief in the efficacy multivitamin against WTP for the multivitamin product, $N = 185$. Error bars indicate standard error.

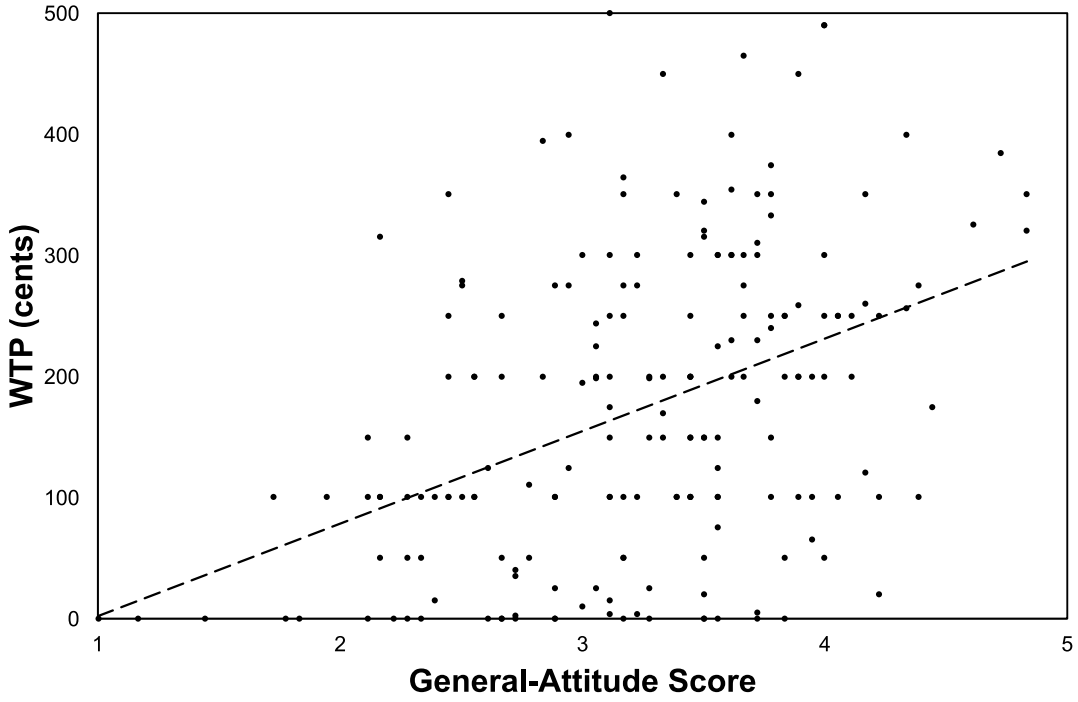

Figure 3.

Participants' general-attitude scale score against WTP-1 for the multivitamin product, N = 194.

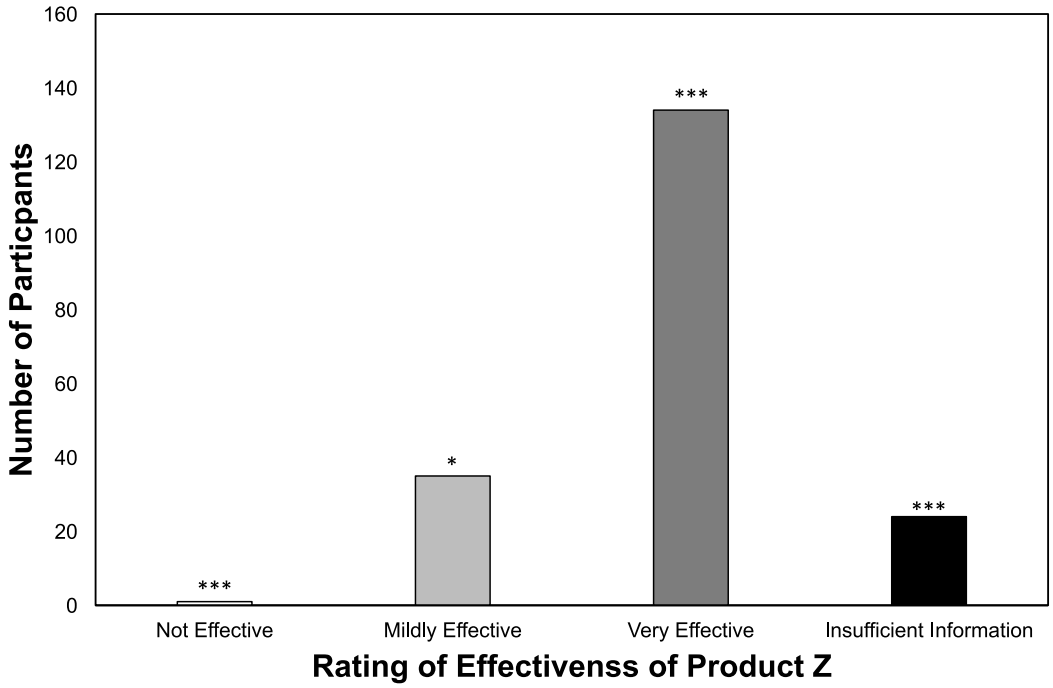

Figure 4.

The number of participants against each possible response concerning the fictitious health product, $N = 194$. Significant differences were determined against the number expected by chance ($N/4$) using planned binomial tests with sequential Bonferroni adjustment.

*$p < .05$, ***$p < .001$

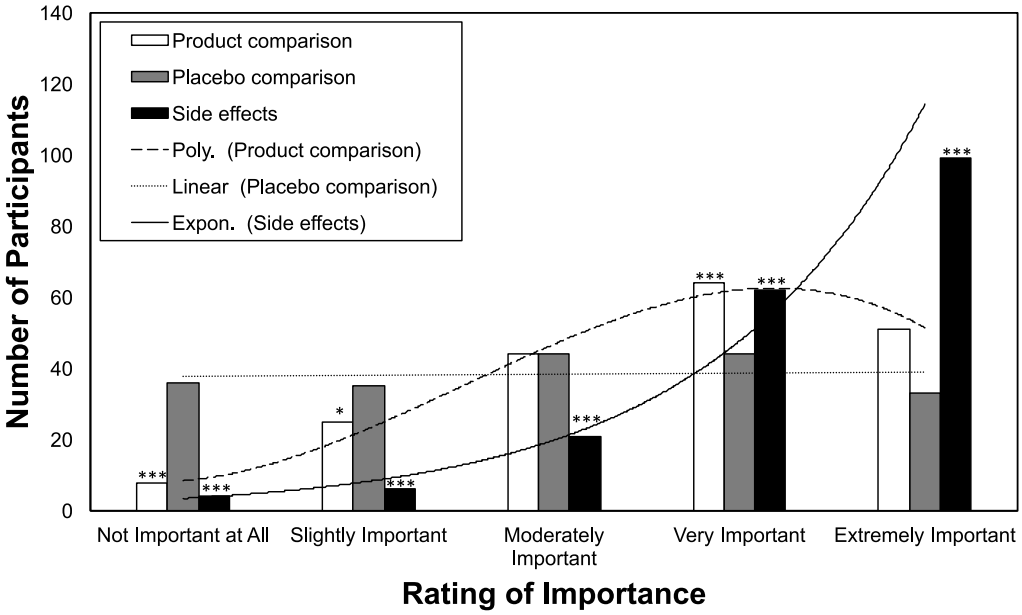

Figure 5.

The number of participants against each possible rating of importance for product comparison, placebo comparison, and side effects, $N = 194$. Significant differences were determined against the number expected by chance ($N/5$) using Wilcoxon sign-ranked tests with sequential Bonferroni adjustment.

*$p < .05$, ***$p < .001$