Working Paper, Jaunary 14, 2021, pp. XX–XX

Multiattribute Judgment: Acceptance of a New COVID-19 Vaccine as a Function of Price, Risk, and Effectiveness

Michael H. Birnbaum*

Abstract

This paper illustrates how to apply the RECIPE design to evaluate multiattribute judgment, reporting an experiment in which participants judged intentions to receive a new vaccine against COVID-19. The attributes varied were Price of the vaccine, Risks of side effects as reported in trials, and Effectiveness of the vaccine in preventing COVID. The RECIPE design is a union of factorial designs in which each of three attributes is presented alone, in pairs with each of the other attributes, and in a complete factorial with all other information. Consistent with previous research with analogous judgment tasks, the additive and relative weight averaging models with constant weights could be rejected in favor of a configural weight averaging model in which the lowest-valued attribute receives additional weight. That is, people are unlikely to accept vaccination if Price is too high, Risk is too high, or Effectiveness is too low. The attribute with the greatest weight was Effectiveness, followed by Risk of side-effects, and Price carried the least weight.

Keywords: COVID, vaccine, Averaging Models, Conjoint Measurement, Functional Measurement, Importance of Variables, Information Integration, Multi-attribute utility, Recipe design, weights of attributes

1 Introduction

In this study, participants were asked to judge their intentions to take a new vaccine against COVID-19, a virus that is highly infectious and which has caused many deaths. At the time of the study, it was a question whether or not people would be willing to take the vaccine, because of disinformation campaigns against vaccination, anti-science dogma, political denials of the dangers of COVID-19, and distrust of the Trump administration in the USA, which had a reputation of promulgating false information. Polls indicated that people might not agree to accept the vaccination in sufficient numbers for a vaccine to produce "herd immunity" and thereby stop the pandemic (Hamel, et al., 2020; Dwyer, 2020).

How would decisions to accept vaccination depend on a vaccine's Price, Risks (of side-effects), and Effectiveness? This topic provides a good illustration of how one can employ the RECIPE design to study multiattribute judgment, using new computer resources that are now available (Birnbaum, 2021).

The Recipe design is an experimental design in which it is possible to distinguish adding and averaging models, and in which weights and scale values in the averaging models can be estimated. The design for three factors consists of the union of each factor alone, the factorial combinations of each pair of factors with the third left out, and the complete factorial design with all three pieces of information. A factorial design does not allow one to distinguish additive from averaging models, nor does it permit any disentanglement of weights from scale values; these facts were presented as criticisms of Anderson's (1974) early work on functional measurement (Schonemann, Cafferty, & Rotton, 1973).

The original RECIPE program was written in FORTRAN as an extension of Birnbaum's (1976) program for a study of intuitive numerical predictions. The Recipe design and program were developed to allow students and researchers how to compare additive and averaging models of information integration (Anderson, 1974) and how to separate weights and scale values (Birnbaum, 1976; Cooke & Mellers, 1998; Mellers & Cooke, 1994; Stevenson, Naylor, & Busemeyer, 1990). Other approaches have also been developed for this issue (Norman, 1976, 1977; Zalinski & Anderson, 1991; Vidotto, Massidda, & Noventa, 2010).

Because few people are still using FORTRAN, Birnbaum (2021) presented three computer programs to update and expand what was previously available. The three programs are Recipe_Wiz.htm, Recipe_sim.htm, and Recipe_fit.xlsx, which enable a user to create Web pages that collect data via the Internet in a Recipe design, simulate data according to a relative weight averaging model with constant weights, and fit data to an averaging model by finding best-fit weights and scale values via the Solver in Excel, respectively. These

Thanks are due to Ulf Reips and Isabel Helm for discussions of their work that inspired this project, and I thank Bonny Quan for useful discussions of issues in this paper.

*Department of Psychology, California State University, Fullerton. Email: mbirnbaum@fullerton.edu.

resources, along with a paper and instructional video that describes them, are available from the following URL: http://psych.fullerton.edu/mbirnbaum/recipe/

The study in this paper was done using these new computer resources. Consistent with previous studies, evidence was found against the relative weight averaging model with constant weights. To fit the data obtained in this study, a new fitting program, Recipe_fit.xlsx, was created to allow for configural weighting. This new resource is included in the Online supplement to this paper, along with the data of this study.

The participants' task was to read descriptions of hypothetical new vaccines for the COVID-19 virus, and to judge their intentions: how likely would they accept a vaccine, based on its Price (P), Risks (R: dangers of side-effects), and Effectiveness (E: how well the vaccine prevented COVID).

1.1 Adding Model

The adding model (Anderson, 1981; Birnbaum & Stegner, 1981; Stevenson, 1993), can be written for this situation as follows:

$$PRE_{ijk} = w_0 s_0 + w_P p_i + w_R r_j + w_E e_k \quad (1)$$

where PRE_{ijk} is the theoretical response in the case where all three attributes, P, R, and E are presented, with levels i, j, and k, respectively, which have scale values of p_i, r_j, and e_k, respectively. The weights (importance) of factors A, B, and C are w_P, w_R, and w_E, respectively. The initial impression has a weight of w_0 and a value of s_0, which represents the response in the absence of information. In the additive model, weights and scale values cannot be separated (Birnbaum & Stegner, 1981; Schoenemann, 1973).

1.2 Averaging Model

The relative-weight averaging model with constant weights (Anderson,1974; 1981; Birnbaum, 1976; Norman, 1976) can be written for this situation as follows:

$$PRE_{ijk} = \frac{w_0 s_0 + w_P p_i + w_R r_j + w_E e_k}{w_0 + w_P + w_R + w_E} \quad (2)$$

where PRE_{ijk} is the theoretical response in the case where all three attributes are presented, with levels i, j, and k, respectively, which have scale values and weights as defined above. The initial impression has a weight of w_0 and a value of s_0. In theory, s_0, represents the value of the impression in the absence of information, and w_0, represents how resistant this "prior" is to new information.

These equations and the following treatment implicitly assume that the judgment function, J, which maps subjective impressions to overt responses, is linear; i.e., that the responses are an interval scale of the subjective impressions.

Key assumptions of these models are: (1) if an attribute is not presented, its weight is zero, (2) the weight of an attribute is independent of the number and value of attributes presented with it. (3) The scale values of attributes are independent of the number and values of the attributes presented with it.

These assumptions imply that there should be no interactions in any of the two-way or three-way factorial designs. Based on weak experiments, Anderson (1974, 1981) argued that the fact that interactions observed in some experiments were minimal or not significant, the failure to reject the null hypothesis "validated" the model, the estimated stimulus scales, and the response scale simultaneously. Both the empirical findings and the logic of the "validation" arguments were disputed by Birnbaum (1982, Section F).

1.3 Configural Weight Averaging Model

Empirical tests of the constant-weight averaging model in better designed studies led to evidence of interactions and other violations of the relative-weight averaging model with constant weights. Evidence of systematic violations led to configural weight theories (Birnbaum, Parducci, & Gifford, 1971; Birnbaum, 1974, 1982, 2008; Birnbaum & Stegner, 1979).

New methods and critical tests were devised to test between the hypotheses that the interactions were due to a nonlinear response function between subjective impressions and over responses, and the hypothesis that the impressions violate the averaging model with constant weights, due to configural weighting. Evidence indicated that one could not explain the violations of the averaging models by means of a nonlinear response function; instead, one needed something like configural weighting (Birnbaum, 1974, 1982, 2008; Birnbaum & Jou, 1990; Birnbaum & Zimmermann, 1998; Birnbaum, 2008).

The configural weight model differs from Equation 2 in that the weight of a given attribute is affected by the configuration of attribute values to be integrated. The range model is the simplest form of configural weight model (Birnbaum, et al., 1971; Birnbaum, 1974; Birnbaum & Stegner, 1979, 1981; Birnbaum & Zimmermann, 1998; Birnbaum, 2018). This model can be written as follows:

$$PRE_{ijk} = \frac{w_0 s_0 + w_P p_i + w_R r_j + w_E e_k}{w_0 + w_P + w_R + w_E} + \omega|max(p_i, r_j, e_k) - min(p_i, r_j, e_k)| \quad (3)$$

where ω is the configural weight transferred from the minimal scale value in the configuration of attribute values, $min(p_i, r_j, e_k)$, to the maximal value, $max(p_i, r_j, e_k)$.

In judgments of morality of a person based on the deeds they have done or the likeableness of a person based on the adjectives that describe them, it has been found that $\omega < 0$.

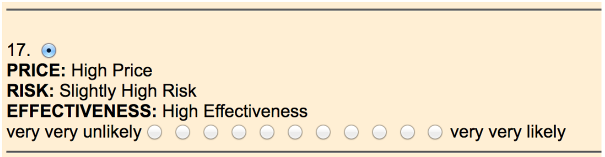

FIGURE 1: Example of display of one trial.

If $\omega < 0$, it means that weight is transferred from the higher-valued information to the lower-valued information. Such configural weighting implies "risk aversion"; if the lower valued outcome in a gamble gets greater weight, people will prefer the the expected value of a gamble to the gamble, even if the utility function is linear (Birnbaum, 2008).

In evaluative judgments, in the "buyer's" point of view, the value of ω is typically negative, but in the "seller's" point of view, it can be positive (Birnbaum & Stegner, 1979; Birnbaum, Coffey, Mellers, & Weiss, 1992; Birnbaum & Sutton, 1992; Birnbaum, et al., 2016; Birnbaum, 2018; Champagne & Stevenson, 1994).

2 Method

This study was done in Fall of 2020, before the FDA had approved a vaccine for COVID-19. Participants viewed the materials by visiting the website and completing a Web form that was created using Recipe_Wiz.htm.

2.1 Instructions

Participants were informed, "COVID-19 virus is a highly contagious disease that can be deadly and can also leave lasting health problems for those who recover from it. Vaccines are currently being developed that are being offered to the public. This questionnaire asks how you would decide whether or not you would take a vaccine based on the price, the risk, and the effectiveness of new vaccines, based on the findings of clinical trials as described by the scientists who conducted the trials."

Price (P) was described as "the amount you must pay out of pocket to receive the vaccine." The 3 levels of P were Low Price: \$20, Medium Price: \$400, and High Price: \$10,000.

Risk (R) was described as ". . . the danger of receiving the vaccine. All medicines and vaccines carry some side effects or risks of bad results that were not intended. The levels of risk are described by the worst outcomes that happened during testing. . ." There were 4 levels of R: Low Risk: 5% of the people got sore arms; Slightly Low Risk: 10% of the people got fevers and headaches for two weeks;Slightly High Risk: 5% of the people got damage to the kidneys;Very High Risk: 1% of those tested had heart attacks.

Effectiveness (E) was described as ". . . how well the vaccine worked to prevent COVID-19 infection in people who were exposed to the virus. The levels of effectiveness are based on the percentage of people who received the vaccine who got sick with COVID-19. . ." The 5 levels of E were: Very Low Effectiveness: 50% got sick; Low Effectiveness: 40% got sick; Medium Effectiveness: 30% got sick; High Effectiveness: 20% got sick; and Very High Effectiveness: 10% got sick.

Each trial was displayed as in the format of Figure 1. Subjects were instructed, "Please make your judgments of whether or not you would be likely to try the vaccine in each case by clicking one of the buttons on the scale from *very very unlikely* to *very very likely* to try the vaccine. In some cases, some of the information is missing, but you should still do the best you can to make your decisions based on the information available."

Complete instructions, warmups, displays, and materials can be found at the following URL:
http://psych.fullerton.edu/mbirnbaum/recipe/vaccine_01.htm

2.2 Design

The Recipe design is based on three factors, designated A, B, and C, with n_A, n_B and n_C levels. It consists of the union of the 3-way factorial design of A by B by C, denoted ABC, combined with each 2-way factorial design (with one piece of information left out), denoted AB, AC, and BC, combined with each piece of information presented alone: A, B, and C. There are a total of $(n_A+1)(n_B+1)(n_C+1)-1$ experimental trials ("cells") in the Recipe design. In this vaccination example, let A = Price (P), B = Risk(R), and C = Effectiveness (E).

There were 3 levels of P, 4 levels of R, and 5 levels of E, producing 119 cells (distinct experimental trials) in the design, consisting of 3 trials for the three levels of Price alone, 4 trials of Risk alone, 5 trials of Effectiveness alone, 3 by 4 = 12 PR trials of Price combined with Risk, 3 by 5 = 15 PE trials of Price by Effectiveness, 4 by 5 = 20 RE trials of Risk by Effectiveness, and 3 by 4 by 5 = 60 PRE trials with all three pieces of information.

2.3 Procedure

These 119 trials were intermixed and presented in random order, following a warm-up of 8 representative trials. Participants were free to work at their own paces, and all completed the task in less than one hour.

2.4 Participants

The participants were 104 college undergraduates who received partial credit (as one option) toward an assignment

TABLE 1: Parameter estimates of configural-weight averaging model.

Price		Risk		Effectiveness	
w_P	0.21	w_R	0.27	w_E	0.38
p_1	1.25	r_1	1.42	e_1	2.07
p_2	5.75	r_2	2.87	e_2	2.41
p_3	9.36	r_3	7.45	e_3	6.35
		r_4	8.86	e_4	10.39
				e_5	10.88

Notes: $w_0 = 0.15$, $s_0 = 5.84$, $\omega = -0.21$. Sum of squared deviations is 11.29; root mean squared error = 0.31.

in Introductory Psychology at California State University, Fullerton. They were tested in Fall of 2020; data collection stopped on December 11, 2020, when the FDA approved the first COVID-19 vaccine in the USA.

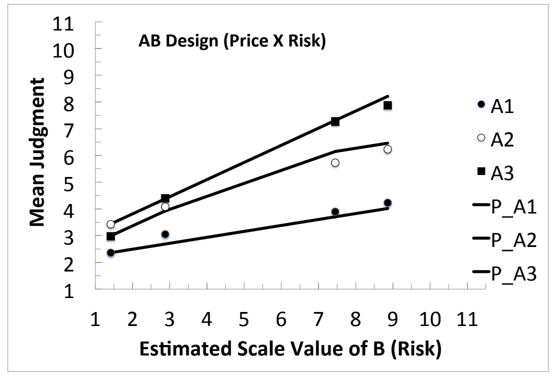

FIGURE 2: Mean judgments of intention to take the new vaccine in the AB design (Price by Risk), as a function of the estimated scale value of Risk (B), with separate markers and curve for each level of Price (A). A1, A2, and A3 refer to Price = \$10,000, \$400, and \$20; the lines show the predictions of configural weight model, labeled P_A1, P_A2, and P_A3, respectively.

3 Results

3.1 Parameter Estimates

The weights and scale values of the configural weight model were estimated to minimize the sum of squared deviations between the mean judgments and the predictions of the model by means of an Excel workbook, Recipe_fit_config.xlsx, which uses the Solver in Excel. This Workbook, including the data of this study, are included in the supplement to this article.

Table 1 shows the best-fit parameters estimated from the data. The weights have been estimated, without loss of generality, such that the sum of the weights is 1. According to the model, $w_E > w_R > w_P$. The configural weight transfer parameter, ω = -0.21, indicating that the lowest-valued attribute receives an additional 21% of the relative weight, transferred from the highest-valued attribute. For example, when all three pieces of information are presented, Effectiveness has a (configural) relative weight of 0.38 – 0.21 = 0.17 when its scale value is highest among the attributes, 0.38 when it is the middle value, and 0.38 + 0.21 = 0.59 when it is the lowest-valued attribute.

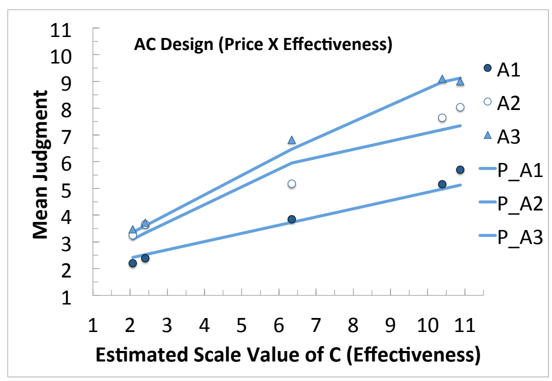

FIGURE 3: Mean judgments in the AC (Price by Effectiveness) design, plotted as a function of estimated scale values of C (Effectiveness), with a separate curve for each level of A (Price); markers show mean judgments and lines show best-fit predictions of the model.

3.2 Two-way Designs

Figures 2, 3, and 4 show mean judgments of intention to accept the new vaccine in the three, two-way factorial subdesigns of Recipe in which one piece of information is missing: AB (Price by Risk), AC (Price by Effectiveness), and BC (Risk by Effectiveness), respectively. In each figure, markers represent mean judgments and lines show best-fit predictions of the configural-weight averaging model (Equation 3).

In Figure 2 mean judgments are plotted against estimated scale values of Risk, with separate markers (and predicted lines) for each level of Price. Both data (markers) and predictions diverge to the right. Such divergence indicates that when either attribute is low in value, the other attribute has less effect.

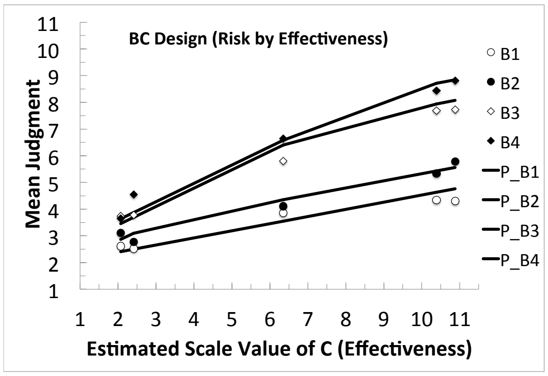

FIGURE 4: Mean judgments in the BC design, plotted as a function of estimated scale values of C (Effectiveness), with separate markers and curve for each level of B (Risk); markers show mean judgments and lines show best-fit predictions of the model.

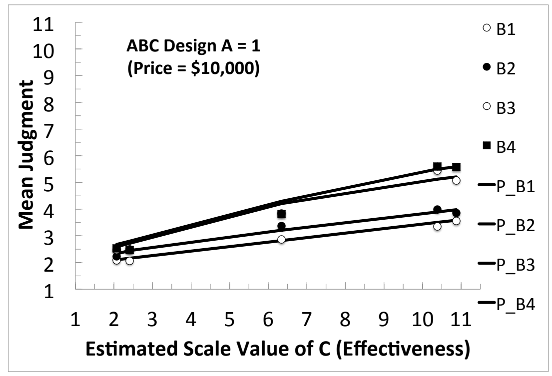

FIGURE 5: Mean judgments of intention to take the new vaccine in the ABC sub-design (Price by Risk by Effectiveness), plotted as a function of the estimated scale value of C (Effectiveness), with a separate curve for each level of B (Risk), where A = 1 (Price = $10,000).

Figure 3 plots mean judgments in the BC sub-design as a function of the estimated scale values of Effectiveness, with a separate curve for each level of price. Figure 4 shows the results for the BC (Risk by Effectiveness) sub-design. In all of the two-way designs, the curves diverge to the right, and the model (lines) does a fairly good job of reproducing the data (markers).

3.3 Three-way Design

Figures 5, 6, and 7 show the mean judgments in the ABC sub-design, which is a 3 by 4 by 5, Price by Risk by Effectiveness, factorial design. Each panel shows the Risk by Effectiveness interaction (plotted as in Figure 4) for a different level of Price, where Price (A) = $10,000, $400, and $20 in Figures 5, 6, and 7, respectively.

There are divergent interactions in the data (markers) for all six cases (Figures 2–7). That is, the vertical separations between the markers increase as one moves from left to right in each figure. These divergent interactions are not consistent with either the additive model or relative weight averaging model with constant weights (Equations 1 and 2). The data (markers) are well-fit by the lines, showing predictions of Equation 3 (the configural weight averaging model), except perhaps in Figure 7 where the divergence in the data is greater than predicted by the model.

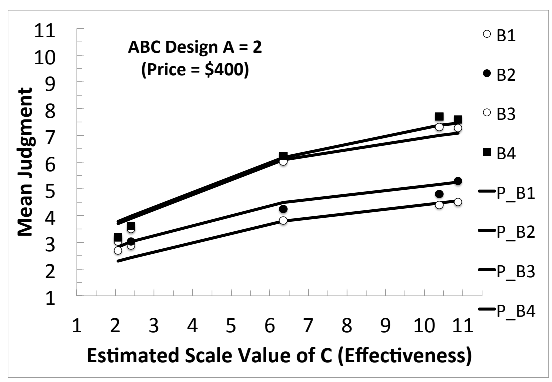

FIGURE 6: Mean judgments in the ABC sub-design (Price by Risk by Effectiveness), as a function of the estimated scale value of Effectiveness, with a separate curve for each level of Risk, where A = 2 (Price = $400).

3.4 Zen of Weights

In the averaging model (Equation 2), the effect of an attribute, like Price or Effectiveness, is directly proportional to the range of scale values multiplied by the weight of a factor, and it is inversely proportional to the sum of the weights of the attributes presented. The term "Zen of Weights" refers to the fact that the effects of A do not inform us clearly about the weight of A, but from the effects of A, we can compare the weights of B and C.

The *effects of A* are defined as differences in response as the factor A is manipulated from A_1 to A_m. Let 1 and m refer to the levels of A that produce the lowest and highest responses for A. The indices, i, j, and k are used for the levels of A, B, and C, respectively, and a bullet ($\bullet$) is used to denote that responses have been averaged over levels of a

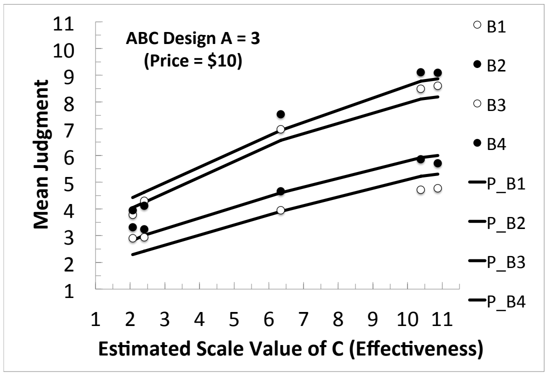

FIGURE 7: Mean judgments in the ABC sub-design (Price by Risk by Effectiveness), as a function of scale values of Effectiveness, with a separate curve for each level of Risk, where Price = $20.

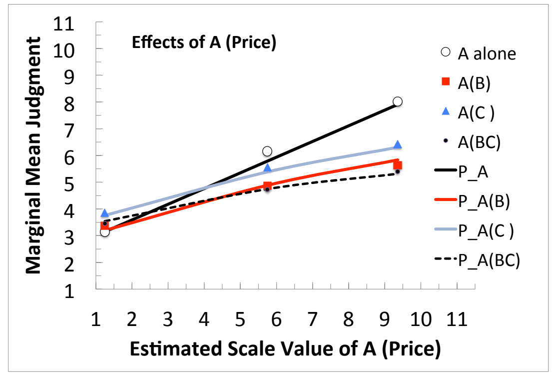

FIGURE 8: Effects of A (Price): Marginal mean judgments as a function of the estimated scale value of Price, with separate markers (data) and curve (predictions) for each sub-design in which A appears.

factor.

When A is presented alone, the effect of A is defined as follows:

$$\Delta A = A_m - A_1 \tag{4}$$

The effects of A in the AB design and AC designs, denoted $\Delta A(B)$ and $\Delta A(C)$, are defined respectively as follows:

$$\Delta A(B) = \overline{AB}_{m\bullet} - \overline{AB}_{1\bullet} \tag{5}$$

$$\Delta A(C) = \overline{AC}_{m\bullet} - \overline{AC}_{1\bullet} \tag{6}$$

where $\overline{AB}_{i\bullet}$ denotes marginal mean in the AB design for level i of A, averaged over the levels of B, and $\overline{AC}_{k\bullet}$ is the corresponding marginal mean for A the AC design, averaged over levels of C.

Finally, the effect of A in the ABC factorial design, denoted $\Delta A(BC)$, is given by,

$$\Delta A(BC) = \overline{ABC}_{m\bullet\bullet} - \overline{ABC}_{1\bullet\bullet} \tag{7}$$

According to the additive model, all of these effects are assumed to be equal; however, according to the relative weight averaging model with constant weights (Equation 2), these effects of A are inversely related to the total weight of the information presented. That is,

$$\Delta A = \Delta a \frac{w_A}{w_0 + w_A} \tag{8}$$

$$\Delta A(B) = \Delta a \frac{w_A}{w_0 + w_A + w_B} \tag{9}$$

$$\Delta A(C) = \Delta a \frac{w_A}{w_0 + w_A + w_C} \tag{10}$$

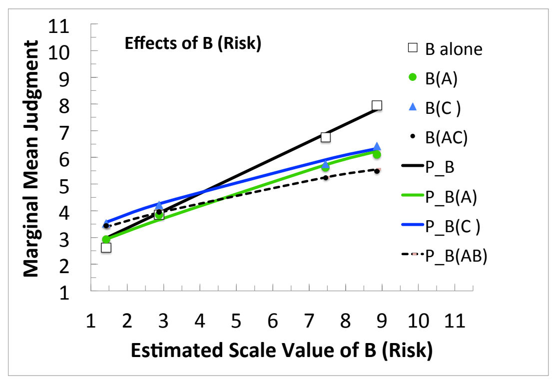

FIGURE 9: Effects of B (Risk): Marginal mean judgments as a function of the estimated scale value of Risk, with separate markers (data) and curve (predictions) for each sub-design including B.

$$\Delta A(BC) = \Delta a \frac{w_A}{w_0 + w_A + w_B + w_C} \tag{11}$$

where $w_A \Delta a$, is the same in all expressions, but the weights in the denominator are different. According to this model, the $\Delta A(BC)$ will be the smallest, and ΔA will be greatest and the other two will be in between such that if the weight of B is greater than the weight of C, then the effect of A will be less when B is presented with it than when it is paired with C.

Figure 8 plots observed and predicted marginal means (according to the configural weight model of Equation 3)

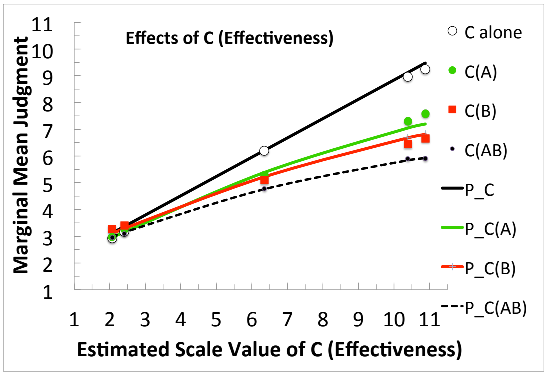

FIGURE 10: Effects of C (Effectiveness): Marginal mean judgments as a function of the estimated scale value of Effectiveness, with separate markers and curve for each sub-design in which C appears.

as a function of the scale values for A; i.e., a_i. Markers represent empirical means or marginal means. Note that the curve for A alone has the steepest slope and the curve for A(BC) has the least slope; that is $\Delta A > \Delta A(BC)$. The fact that these slopes (effects) are not equal (the curves even cross) is evidence against the adding model, but consistent with either form of averaging (Equations 2 or 3).

Figures 9 and 10 show effects of B and C, as in Figure 8. Figure 9, shows that the effect of B(A), Risk averaged over levels of Price (slope of the solid circles) exceeds the effect of B(C), Risk averaged over Effectiveness; therefore the weight of Effectiveness is greater than that of Price. In Figure 10, the effect of C(A) exceeds that of C(B); therefore, the weight of A (Price) is the least weighted attribute.

4 Discussion

The data allow us to reach two main conclusions: First, judgments of intention to use a vaccine do not conform to an additive model (as in Equation 1), but instead to an averaging model (as in Equations 2 or 3). The fact that the slopes in Figures 8, 9, and 10 differ from each other rules out the adding model, but the finding that slopes decrease as more information is presented would be compatible with either Equations 2 or 3.

Second, the data do not conform to the predictions of the relative weight averaging model with constant weights. This model implies that the data in Figures 2–7 should be parallel. The fact that the data in Figures 2–7 show divergent interactions violates the implications of that model. Instead, the data can be better fit by an averaging model with configural weights (Equation 3), in which the lower-ranked information receives greater weight.

The data also indicate that Price is the least important factor and that Effectiveness is the most important factor in deciding whether to accept vaccination. Of course, the weights and scale values in Table 1 represent average behavior of college undergraduates. It would not be surprising if other populations, including older people who might be at greater risk for consequences of COVID-19, might have different relative weights for risks of the vaccine against risks of the disease from those of college students.

Although this is the first study (to my knowledge) of vaccine acceptance using a Recipe design, I would argue that the two main findings could have been "predicted" by generalization of findings from previous research. The concepts of "generalization" and "similarity" will be discussed in more detail in the next section.

First, Research with similar tasks concluded that additive models can be rejected in favor of averaging models because the effect of a component has been found to be inversely related to the number and importance of components with which it is combined (Anderson, 1974; 1981; Birnbaum, 1976; Birnbaum, Wong, & Wong, 1976; Birnbaum & Mellers, 1983).

Second, research with evaluative judgments concluded that the relative weight averaging model with constant weights (Equation 2), advocated by Anderson (1974, 1981), can be rejected in favor of configural weighting models (Equation 3). Judgments in similar studies show divergent interactions like those observed here in Figures 2–7 (Birnbaum, 1972, 1973, 1974).

Empirically, many studies have observed interactions that have been interpreted as evidence of configural weighting (Birnbaum, 1974; 2008; Birnbaum & Stegner, 1979, 1981; Birnbaum & Zimmermann, 1998; Champagne & Stevenson, 1994). Although these violations of the constant-weight averaging model refute the simple constant-weight averaging model (that implies parallel lines in Figure 2), they appear compatible with a configural weight averaging model that has the same implications regarding the relative effects of variables in the recipe design (the slopes in Figure 1).

Because interactions could be induced by a nonlinear transformation between subjective impressions and overt responses, research was conducted to compare two theories: (1) Equation 2 with a nonlinear judgment function versus (2) Equation 3 with a linear judgment function. This research required introduction of new experimental designs and techniques involving scale convergence, scale-free tests, testing joint independence, and models of judgment and reaction times. The findings of the research led to a coherent conclusion that the divergent interactions obtained in evaluative judgments could not be explained away by a nonlinear judgment function, but instead, configural weighting could be retained as the explanation of these various studies (Birnbaum, 1974; 1982; 2008, 2018; Birnbaum & Jou, 1990;

Birnbaum & Zimmermann, 1998).

4.1 Similar Tasks: Analogies

What is the basis for saying that one can generalize from one experiment to another that is similar? The basis for generalization is the theory that common mechanisms for intuitive aggregation of information are induced by these tasks. Deciding to accept a new vaccine is similar to judging the likeableness of a person described by a set of adjectives or judging the morality of a person based on their deeds. Table 2 presents a table of tasks that have reported divergent interactions. All involve evaluative judgments based on attributes, cues, or components that convey evaluation. (See Birnbaum and Mellers (1983, Table 2) for another table like this, used to organize other experiments, including studies that do not observe prominent interactions.)

Birnbaum, Wong, and Wong (1976) asked people to judge how likeable is a person described by a set of adjectives contributed by people who had different lengths of acquaintance with the person described. For example, how much do you think you would like a person who was described by an acquaintance of 3 years as "kind" and by a person who met the person only once as "phony"? To represent the judgments, scale values differed for different adjectives, and weights depended on the length of acquaintance with the target person. As in Birnbaum (1974), there is a significant divergent interaction such that even when the sources are equal in length of acquaintance, the lower-valued adjective appears to receive greater weight.

In Moral judgments, a person who has done a very bad deed, such as *killing one's mother without justification*, is rated as "immoral" even if that person has done a number of good deeds such as *donating a kidney to a child needing a transplant* (Birnbaum, 1972, 1973; Riskey & Birnbaum, 1974). Although the more good deeds a person had done the higher the judgment, it appears that a person's worst deed sets an upper limit on how moral that person can be judged.

Birnbaum and Stegner (1979) asked people to judge the most a buyer should be willing to pay for a used car, based on estimates from sources that varied in both bias and expertise. The sources were mechanics who had examined the cars who varied in their mechanical expertise and their relationships to the potential buyer and seller of the car (they were friends of the buyer, seller, or independents). It was concluded that the lower valued estimates received greater configural weight in determining buying prices. Birnbaum and Zimmermann (1998) found similar results for buying prices of investments.

Divergent interactions have also been observed in judgments of the buying prices of gambles based on the possible cash outcomes of the gambles (Birnbaum & Sutton, 1992; Birnbaum, Coffey, Mellers, & Weiss, 1992; Birnbaum, et al., 2016). Birnbaum, Thompson, & Bean (1997) found violations of interval independence, which agreed in direction with the interactions. Studies of branch independence (aka, joint independence) in gambles and investments have also confirmed that the interactions in buying prices of gambles and investments are "real" and cannot be explained by nonlinear judgment functions (Birnbaum & Beeghley, 1997; Birnbaum & Veira, 1998; Birnbaum & Zimmermann, 1998). Violations of restricted branch independence in choices between gambles shows that configural weighting remains the best explanation of these interactions (Birnbaum, 2008; Birnbaum & McIntosh, 1996).

Champagne and Stevenson (1994) asked people to combine information about an employee's job performance for the purpose of rewarding good performance. They reported similar divergent interactions, in which poor performance in one component diminished the effect of other components.

4.2 Judge's Point of View

There is evidence that the divergent interactions can be reduced and even reversed by changing the participant's point of view (Birnbaum & Stegner, 1979). If instead of asking people to judge the highest price that a buyer should be willing to pay to buy something, we ask people to judge the least that a seller should be willing to accept, then we can eliminate or reverse the interaction (Birnbaum & Stegner, 1979; Birnbaum & Beeghley, 1997; Birnbaum & Sutton, 1992; Birnbaum, et al., 1992; Birnbaum, et al., 2016; Birnbaum, 2018).

Whereas Birnbaum and Stegner (1979) had described the difference between willingness to pay and willingness to accept in terms of the judge's point of view and configural weighting, others referred to this phenomenon as the "endowment" effect, and tried to explain it with extensions of prospect theory (Thaler, 1980; Schmidt, Starmer, & Sugden, 2008), Birnbaum and Zimmermann (1998), Birnbaum, et al. (2016), and Birnbaum (2018) show how that rival theory cannot account for the data because buying and selling prices are not monotonically related to each other and because buying and selling prices violate complementary symmetry.

In moral judgments or likeableness judgments, the judge is in the viewpoint of deciding whether to accept another person, as in the buyer's point of view. The seller's point of view occurs when a person states how he or she should be judged. Students ask to be judged by their best performance, and ask that their lowest exam score be dropped. I suspect that people would like to be judged in likeableness or morality on the basis of their best traits or deeds, and ask that their worst qualities or deeds be forgiven.

Champagne and Stevenson (1994) asked participants to judge job performance not only for the purpose of rewards, but also for the purpose of possible punishments. In the case of punishments, the interaction was of the opposite direction; that is, good performance components received greater weight. This finding also fits with the concept that the

TABLE 2: Analogies among studies of judgment showing divergent interactions.

Judgment	Weight	Scale Value
Vaccination	attributes importance	levels of attributes
Likeableness	length of acquaintance of source	adjectives
Morality	number of deeds	deed types
Buying price used car	source expertise	source's estimate
Buying price of gambles	probability to win	prize values
Employee reward	components of job performance	levels of performance

See text for descriptions of these studies.

judge's point of view can affect the magnitude and direction of the configural weighting.

There are also situations that are (partially) analogous but in which interactions are not prominent, such as intuitive numerical predictions. Birnbaum (1976) asked people to predict a numerical criterion, based on independent numerical cues that are correlated with the criterion. In this case, interactions were not prominent, and the relative weight averaging model of Equation 2 achieves a satisfactory fit, with weights that are greater for cues that are more highly correlated with the criterion.

Taking this summary of previous research into consideration, one can make the following prediction for new tasks, such as multiattribute consumer judgments of products (Meyer, 1981; Mellers & Cooke, 1994): From the buyer-consumer's perspective, multiattribute products are predicted to show divergent interactions between attributes, such that if an important attribute is low in value, the item will be judged low in value, and other attributes of a product may improve the judgment but by less than if that important attribute were high in value.

4.3 Vaccination Disinformation

Aside from academic questions comparing theoretical models of how judges combine information when they are placed in various points of view, there are practical issues in the real world. For example, how should a reasonable person communicate when sources of information are not only biased, but may have incentives to lie. In politics, legal settings, and intelligence (e.g., spying), there can be tactical advantages to using disinformation.

A fuller discussion of political persuasion, indoctrination, faith, and cults are beyond the scope of this paper, but I think it worthwhile to mention a "real world" issue that will affect acceptance of a vaccine, namely, disinformation campaigns from the "anti-vaxxer" movement, perhaps financed by foreign intelligence (e.g., DiResta, 2020). In order for a vaccine to be effective in stopping a pandemic, a certain percentage of people must be convinced to take the vaccine. If the media provides equal time to scientists and to equally persuasive "anti-vaxxers", then from the averaging model, one would expect that about half the public would decide that vaccines provide more benefits than risks. Even if 99% of scientists say a vaccine is safe and effective and 1% says the risk of harm outweighs the benefits, if the media presents one representative of each "side" in the "debate", the public would be predicted to split 50-50, based on Equation 2, and based on the configural weight model with the value of $\omega = -0.2$, the average acceptance of the vaccine would be predicted to be 30%.

The situation becomes even more complicated if people live in different media "bubbles" in which only one of the two sides presents most of the messages. Such a situation can produce a society of people with strong, but opposite convictions.

References

Anderson, N. H. (1974). Information integration theory: A brief survey. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & D. Suppes (Eds.), *Contemporary developments in mathematical psychology*, vol. 2. San Francisco: W. H. Freeman.

Anderson, N. H. (1981). *Foundation of information integration theory*. New York: Academic Press.

Birnbaum, M. H. (1972). Morality judgments: Tests of an averaging model. *Journal of Experimental Psychology, 93*(1), 35-42.

Birnbaum, M. H. (1973). Morality judgment: Test of an averaging model with differential weights. *Journal of Experimental Psychology, 99*(3), 395-399.

Birnbaum, M. H. (1974). The nonadditivity of personality impressions. *Journal of Experimental Psychology Monograph, 102*, 543-561.

Birnbaum, M. H. (1976). Intuitive numerical prediction. *American Journal of Psychology, 89*, 417-429.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp 401-485). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review, 115*, 463-501.

Birnbaum, M. H. (2018). Empirical evaluation of third-generation prospect theory. *Theory and Decision, 84*(1), 11-27.

Birnbaum, M. H. (2021). Computer resources for implementing the recipe design for weights and scale values in multiattribute judgment. *Submitted for Publication, xx*, xx-xx. WWW Document, viewed January 2, 2021, http://psych.fullerton.edu/mbirnbaum/recipe/

Birnbaum, M. H., & Beeghley, D. (1997). Violations of branch independence in judgments of the value of gambles. *Psychological Science, 8*(2), 87-94.

Birnbaum, M. H., Coffey, G., Mellers, B. A., & Weiss, R. (1992). Utility measurement: Configural-weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 331-346.

Birnbaum, M. H., & Jou, J. W. (1990). A theory of comparative response times and "difference" judgments. *Cognitive Psychology, 22*(2), 184-210.

Birnbaum, M. H., & McIntosh, W. R. (1996). Violations of branch independence in choices between gambles. *Organizational Behavior and Human Decision Processes, 67*(1), 91- 110.

Birnbaum, M. H., & Mellers, B. A. (1983). Bayesian inference: Combining base rates with opinions of sources who vary in credibility. *Journal of Personality and Social Psychology, 45*, 792-804.

Birnbaum, M. H., Parducci, A., & Gifford, R. K. (1971). Contextual effects in information integration. *Journal of Experimental Psychology, 88*(2), 158-170.

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology, 37*, 48-74.

Birnbaum, M. H., & Stegner, S. E. (1981). Measuring the importance of cues in judgment for individuals: Subjective theories of IQ as a function of heredity and environment. *Journal of Experimental Social Psychology, 17*, 159-182.

Birnbaum, M. H., & Sutton, S. E. (1992). Scale convergence and utility measurement. *Organizational Behavior and Human Decision Processes, 52*(2), 183-215.

Birnbaum, M. H., Thompson, L. A., & Bean, D. J. (1997). Testing interval independence versus configural weighting using judgments of strength of preference. *Journal of Experimental Psychology: Human Perception and Performance, 23*(4), 939-947.

Birnbaum, M. H., & Veira, R. (1998). Configural weighting in judgments of two- and four-outcome gambles. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 216-226.

Birnbaum, M. H., Wong, R., & Wong, L. (1976). Combining information from sources that vary in credibility. *Memory & Cognition, 4*(3), 330-336.

Birnbaum, M. H., Yeary, S., Luce, R. D., & Zhao, L. (2016). Empirical evaluation of four models for buying and selling prices of gambles. *Journal of Mathematical Psychology, 75*, 183-193.

Birnbaum, M. H., & Zimmermann, J. M. (1998). Buying and selling prices of investments: Configural weight model of interactions predicts violations of joint independence. *Organizational Behavior and Human Decision Processes, 74*(2), 145-187.

Champagne, M., & Stevenson, M. K. (1994). Contrasting models of appraisal judgments for positive and negative purposes using policy modeling. *Organizational Behavior and Human Decision Processes, 59*, 93-123.

Cooke, A. D. J., & Mellers, B. A. (1998). Multiattribute judgment: Attribute spacing influences single attributes. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 496–504.

DiResta, R. (2020). Anti-vaxxers think this Is their moment. *The Atlantic.* WWW document, viewed January 7, 2021. https://www.theatlantic.com/ideas/archive/2020/12/campaign-against-vaccines-already-under-way/617443/

Dwyer, C. (2020). Poll: Americans are growing less reluctant to take COVID-19 vaccine. WWW Document, Viewed January 2, 2021.

Hamel, L., Lopes, L., Muñana, C., Artiga, S., & Brodie, M. (2020). KFF/The Undefeated Survey on Race and Health. WWW Document, viewed January 2, 2021. https://www.kff.org/racial-equity-and-health-policy/report/kff-the-undefeated-survey-on-race-and-health/

Mellers, B.A., & Cooke, A. D. J. (1994). Tradeoffs Depend on Attribute Range. *Journal of Experimental Psychology: Human Perception and Performance 20*, 1055–1067.

Meyer, R. J. (1981). A model of multiattribute judgments under attribute uncertainty and informational constraint *Journal of Marketing Research, 18*(4), 428-441.

Norman, K. L. (1976). A solution for weights and scale values in functional measurement. *Psychological Review, 83*, 80-84. DOI: 10.1037/0033-295X.83.1.80

Norman, K.L. (1977). Attributes in Bus Transportation: Importance Depends on Trip Purpose. *Journal of Applied Psychology 62*, 164–170.

Riskey, D. R., & Birnbaum, M. H. (1974). Compensatory effects in moral judgment: Two rights don't make up for a wrong. *Journal of Experimental Psychology, 103*(1), 171-173.

Schmidt, U., Starmer, C., & Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty,36*, 203–223.

Schonemann, P. H., Cafferty, T., & Rotton, J. (1973). A note on additive functional measurement. *Psychological Review, 80*(1), 85–87.

Stevenson, M. K. (1993). Decision making with long-term consequences: Temporal discounting for single and multiple outcomes in the future. *Journal of Experimental Psychology: General, 122*(1), 3-22.

Stevenson, M. K., Busemeyer, J. R., & Naylor, J. C. (1990). Judgment and decision-making theory. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (p. 283–374). Consulting Psychologists Press.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization, 1*, 39–60.

Vidotto G., Massidda D., & Noventa S. (2010). Averaging models: parameters estimation with the R-Average procedure *Psicológica, 31*,461-475.

Zalinski, J., & Anderson, N. H. (1991). Parameter estimation for averaging theory. In N. H. Anderson (Ed.), *Contributions to Information Integration Theory* (Vol. 1, pp. 353–394). Hillsdale, NJ: Lawrence Erlbaum Associates.