Psychological Methods

Modeling Multiple Response Processes in Judgment and Choice

Ulf Böckenholt

Online First Publication, April 30, 2012. doi: 10.1037/a0028111

Psychological Methods

1082-989X/12/$12.00 DOI: 10.1037/a0028111

Modeling Multiple Response Processes in Judgment and Choice

Ulf Böckenholt
Northwestern University

In this article, I show how item response models can be used to capture multiple response processes in psychological applications. Intuitive and analytical responses, agree–disagree answers, response refusals, socially desirable responding, differential item functioning, and choices among multiple options are considered. In each of these cases, I show that the response processes can be measured via pseudoitems derived from the observed responses. The estimation of these models via standard software programs that allow for missing data is also discussed. The article concludes with two detailed applications that illustrate the prevalence of multiple response processes.

Keywords: item response models, missing data, multiple-choice items

Supplemental materials: http://dx.doi.org/10.1037/a0028111.supp

A key challenge in quantitative psychology is to develop models that parsimoniously capture how individuals differ in arriving at their judgments or choices. Successful examples include item response and discrete choice models (Böckenholt, 2006; van der Linden & Hambleton, 1997). Both classes of models have in common that they postulate a single response process that leads to the observed judgments and choices. In item response models, the probability of a correct response to an item depends on the difference between the test taker's ability and the item difficulty. In choice models, the probability of a choice depends on the differences in utility between the choice options. Thus, in both cases, a single response process, formalized as a difference between ability and difficulty for item response models or as a difference between utilities for choice models, is postulated to hold for all respondents. In this article, I go beyond the notion of a single response process and consider applications in which respondents may arrive at their answers via multiple response processes.

Multiple response processes abound in psychological research. For example, there are a considerable number of dual-response theories in judgment and choice applications that are based on System 1 and System 2 distinctions (Evans, 2008). System 1 processes are characterized as unconscious, rapid, effortless, and automatic, whereas System 2 processes are characterized as conscious, slow, effortful, and deliberative. Each system can lead to different answers, as illustrated by the following test item: "A bat and a ball cost $1.10. The bat costs $1 more than the ball. How much does the ball cost?" (Frederick, 2005, p. 26). A typical immediate answer is "10 cents" because $1.10 can be divided easily into $1 and 10 cents, and 10 cents seems to be a reasonable price for a ball. However, after a moment of reflection and deliberation, a respondent may realize that the difference between $1 and 10 cents is less than $1 and give the correct answer instead.

Similarly, when asked questions about personal or sensitive issues, respondents may want to give honest answers but also want to present themselves in a favorable light, with the result that items measure both the actual behaviors of the respondents as well as the respondents' tendency to edit their responses. To identify which response process gives rise to the observed answer, social desirability scales (Paulhus, 1984) have been developed that measure the degree to which respondents tend to present themselves favorably. However, success in using these scales to correct for respondents' response-editing behavior has been limited so far (Steenkamp, De Jong, & Baumgartner, 2010).

In a choice task, decision makers may differ in their sensitivity to contextual cues (e.g., whether an option is a compromise among the available options) and their preferences for the presented attributes of the choice options. Some decision makers may base their choice on the basis of contextual cues only, whereas others may focus on the observed attributes of the choice options. Systematic violations of random utility theories are observed when these different ways of choosing among the available options are not taken into account (Tsetsos, Usher, & Chater, 2010).

Allowing for multiple response processes may also lead to more valid conclusions about both item characteristics and drivers of individual differences. For example, in the previously mentioned dual-process application, experimental treatments (e.g., the readability of the choice material; Alter, Oppenheimer, Epley, & Eyre, 2007) may affect System 1 but not System 2 processes. Survey methods (e.g., whether the interview is conducted face to face or via an online questionnaire) may affect the degree to which respondents edit their responses but not the actual behavior under study. Also, motivational states may moderate attention to contextual cues but not the attention toward the options' attributes (Mourali, Böckenholt, & Laroche, 2007). Thus, in all of these cases, relating covariates to the hypothesized response processes or direct experimental manipulations of the response subprocesses

Ulf Böckenholt, Kellogg School of Management, Northwestern University.

I am grateful to Carolyn Roux and Jacques Nantel for the rating data presented in the Applications section of this article.

Correspondence concerning this article should be addressed to Ulf Böckenholt, 2001 Sheridan Road, Evanston, IL 60208. E-mail: u-bockenholt@kellogg.northwestern.edu

can prove critical for understanding the determinants of the responses.

In this article, I focus on multiple response processes in judgment and choice applications that can be represented via a tree structure. The emphasis on nested response processes is motivated not only by the reported applications but also by the consideration that this class of multiple response models can be fit with standard software programs for item response models that allow for missing data. Specifically, I show that Mplus (Muthén & Muthén, 1998–2010) is well suited for estimating tree-based response models with process-specific covariates.

The proposed approach belongs to the class of latent response models (Maris, 1995) because the observed response is viewed as a result of multiple latent responses. Related models include ordinal step models (Verhelst, Glas, & Vries, 1997; Tutz, 1997), which were developed originally as an alternative to the partial credit model (Masters, 1982) for the analysis of educational testing data. Here I go beyond this original application and show that both ordinal and nominal responses can be decomposed to test for the presence of multiple response processes. I also discuss links to differential item functioning, multinomial processing tree models (Batchelder, 2009), and diagnostic measurement models (Rupp, Templin, & Henson, 2010).

The next section introduces the tree structure designed to capture multiple response processes and shows how the branches of a tree can be represented by pseudoitems. Subsequently, two applications demonstrate the usefulness of this approach in both judgment and choice settings. The article concludes with a discussion of future avenues for research.

Multiple-Response-Process Models

This section presents the tree-structure framework for modeling multiple response processes. I first present several examples that illustrate the relationships between the observed response and the postulated latent response processes. Next, I show how the decomposition of the observed response into nested response subprocesses is facilitated by the use of pseudoitems. I also discuss the inclusion of covariates that are specific to each of the latent response processes. To illustrate the versatility of this approach, the subsequent section presents analyses of Likert responses and choice data.

Multiple Response Processes

Assume that there are J polytomous items that are either nominal or ordinal in nature. For each of these items, we postulate that the observed response is a result of multiple latent response processes. Figures 1, 2, and 3 depict increasingly more complex tree structures representing the relation between the observed and latent response processes. Specifically, Figure 1 illustrates the standard item response representation for a binary item with the two categories, A and B. We assume that the response process can be represented by the one-parameter probit model (Birnbaum, 1968), which allows both person and item effects to be estimated. Under this model, the probability that person i selects item j's response category A can be written as

$$\Pr(y_{ij} = A) = 1 - \Phi(\theta_i - \gamma_j), \qquad (1)$$

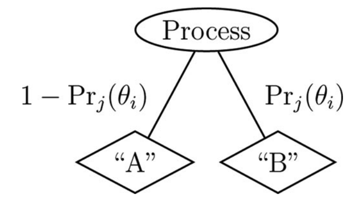

Figure 1. Tree diagram of the single-process model. This figure represents the probabilistic outcomes of a binary item. The probabilities of observing Outcomes A and B are given by the respective branch probabilities. The process is unobserved and assumed to be captured by the one-parameter probit model.

where Φ is the normal cumulative distribution function and γ_j represents the location of item j on the latent continuum defined by θ. Correspondingly, the probability that response category B is selected is given by the complement of Equation 1 with

$$\Pr(y_{ij} = B) = \Phi(\theta_i - \gamma_j). \qquad (2)$$

In both equations, the person-specific effect is represented by the parameter θ_i, which is assumed to follow a normal distribution in the population of respondents. The branch probabilities in Figure 1 are given by Equation 1 with $\Pr(y_{ij} = A) = 1 - \Pr_j(\theta_i)$ and by Equation 2 with $\Pr(y_{ij} = B) = \Pr_j(\theta_i)$.

Responses to a trichotomous item may be a result of two latent processes that are represented by two subtrees, as depicted in Figure 2. Here the first response process (denoted by *I*) yields the outcome A with probability $\Pr_j(\theta_i^{(I)})$. With the complementary probability $1 - \Pr_j(\theta_i^{(I)})$ the second response process is activated, which yields the outcome B with probability $\Pr_j(\theta_i^{(II)})$ and C with complementary probability $1 - \Pr_j(\theta_i^{(II)})$. Note that there are two person-specific parameters, $\theta_i^{(\mathrm{I})}$ and $\theta_i^{(\mathrm{II})}$, that describe individual differences captured by the two pseudoitems. The Roman letters *I* and *II* are used to emphasize the order of the two subprocesses. Thus, B or C can be observed only after the first response process is concluded.

The notion of two nested subprocesses captures the intuitive-deliberate distinction described in the introduction. When contemplating the question about the cost for a ball, a respondent may arrive initially at the intuitive response of 10 cents on the basis of System 1 processing. This response is the final answer if the respondent does not have sufficient inhibitory control to engage in systematic, effortful thought to override this response. However, if inhibitory control is available to suppress the System 1 response, deliberate processing can revise the intuitive response and arrive at the correct answer, provided no algorithmic or other calculation errors are committed. Thus, in this case, A refers to the intuitive response, B refers to incorrect responses based on deliberate reasoning, and C refers to the correct response. The first response process (I) captures ability differences in inhibitory control (Logan, Schachar, & Tannock, 1997) and the second response process (II) captures ability differences in deliberate reasoning (Böckenholt, 2012).

In the Applications section of this article, I apply a similar set-up to disentangle the effects of contextual and attribute-based information in choice. Here the three response categories correspond to different choice options that vary on two negatively correlated attributes (e.g., price and quality). Options A and C score high on

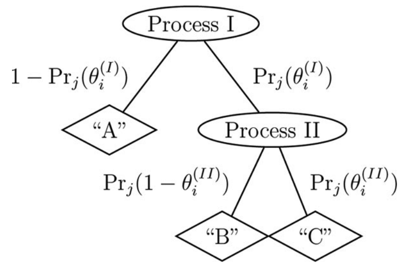

Figure 2. Tree diagram of the two-process model. This figure represents the probabilistic outcomes of a trichotomous item. The probabilities of observing Outcomes A, B, and C are given by the product of the respective branch probabilities. Both Processes I and II are unobserved and assumed to be captured by the one-parameter probit model.

one of the two attributes, whereas Option B can be viewed as a compromise by taking on intermediate values on both attributes. The tree model in Figure 2 illustrates that decision makers may first consider contextual information and decide whether to pick the compromise option, B. If the compromise is not picked, then, as a next step, they choose between Options A and C on the basis of the preferred attribute.

Extensions to items with more than three response categories allow for more complex tree structures. Figure 3 shows a three-process scenario for an item with five response categories. Process I leads either to the selection of Response Category C or to the activation of Process II, which, in turn, activates Process III. The latter process yields the outcomes A and B or D and E. In the Applications section, I report an example for such subprocesses that captures how a person may answer an attitudinal item with the response categories *strongly disagree*, *disagree*, *neither disagree or agree*, *agree*, and *strongly agree*. Here we assume that respondents first decide whether to express an opinion or not express an opinion (which corresponds to the middle category, C). At the next steps, we assume that respondents decide on the direction of their attitudes and, finally, on the intensity of their attitudes.

In sum, each polytomous item can be represented by multiple nested processes that give rise to the selection of one of the response categories. To facilitate the estimation of multiple response models, I introduce the restriction that each observed response category has a unique path to one of the latent response processes. Thus, scenarios under which the same response category may be arrived at via different latent processes are not considered. For example, this restriction excludes testing applications in which respondents may give a correct answer to a test item on the basis of reasoning or guessing. Although I acknowledge that because of this limitation, I do not provide a complete solution to the modeling of multiple latent response processes, it is worth stressing that the considered class of latent response models is of much use in applied work.

Pseudoitems

For the estimation of the multiple response models, the outcomes of each response process are represented by pseudoitems. For example, for the tree presented in Figure 2, two pseudoitems, I and II, are introduced that correspond to the two latent processes. Table 1 summarizes the outcomes of the two pseudoitems and the corresponding selections of the response categories. Specifically, Response A is obtained when the first pseudoitem takes on the value 0 and a missing value by design is obtained for the second pseudoitem. Response B is observed when the first and second pseudoitems take on the values 1 and 0, respectively. And Response C is obtained when both the first and second pseudoitems take on the value 1. Assuming that the one-parameter probit model (Birnbaum, 1968) can be used to represent the response probabilities for both pseudoitems, we obtain for the three response categories

$$\Pr(y_{ij} = A) = 1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)}), \tag{3}$$

$$\Pr(y_{ij} = B) = \Phi(\theta_i^{(I)} - \gamma_j^{(I)})(1 - \Phi(\theta_i^{(II)} - \gamma_j^{(II)})), \tag{4}$$

and

$$\Pr(y_{ij} = C) = \Phi(\theta_i^{(I)} - \gamma_j^{(I)})\Phi(\theta_i^{(II)} - \gamma_j^{(II)}), \tag{5}$$

where $\gamma_j^{(I)}$ and $\gamma_j^{(II)}$ are the respective item parameters for the two sequential processes, and $\theta_i^{(I)}$ and $\theta_i^{(II)}$ represent the respective person parameters. These two person parameters are assumed to

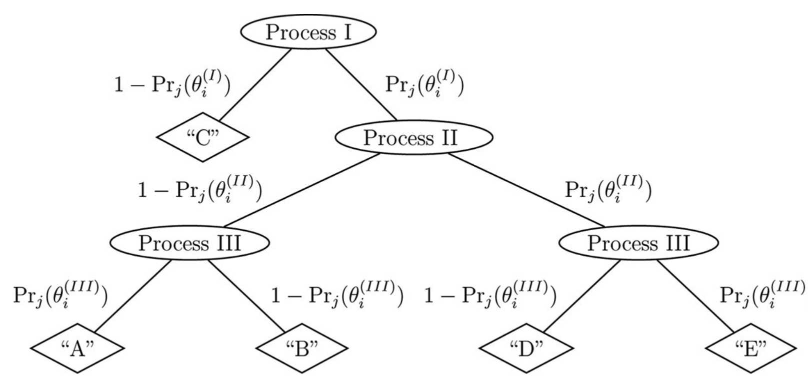

Figure 3. Tree diagram of the three-step model. This figure represents the probabilistic outcomes of a five-category item. The probabilities of observing the Outcomes A–E are given by product of the respective branch probabilities. Processes I, II, and III are unobserved and assumed to be captured by the one-parameter probit model.

Table 1
Pseudoitems of the Two-Process Model

Category	Item I	Item II	Category probabilities
A	0	—	$1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)})$
B	1	0	$\Phi(\theta_i^{(I)} - \gamma_j^{(I)})\,[1 - \Phi(\theta_i^{(II)} - \gamma_j^{(II)})]$
C	1	1	$\Phi(\theta_i^{(I)} - \gamma_j^{(I)})\,\Phi(\theta_i^{(II)} - \gamma_j^{(II)})$

Note. The dash indicates data are missing by design.

follow a bivariate normal distribution, with means 0, standard deviations $\sigma^{(I)}$ and $\sigma^{(II)}$, and covariance $\sigma^{(I,II)}$.

Similarly, for the tree model depicted in Figure 3, the outcomes of the subprocesses can be represented by three pseudoitems as shown in Table 2. Here Item I takes on the value 0 when Category C is selected and the value 1 otherwise. Item II takes on the value 0 when Categories A and B are selected and the value 1 when Categories D and E are selected. Item III takes on the value 1 when either Category B or Category D is selected and the value 0 for Categories A and E. By design, values for Items II and III are missing for Response Category C. Using the one-parameter probit model (Birnbaum, 1968) to represent the probabilities for the three pseudoitems, we obtain for the five item categories

$$\Pr(y_{ij} = A) = \Phi(\theta_i^{(I)} - \gamma_j^{(I)})[1 - \Phi(\theta_i^{(II)} - \gamma_j^{(II)})] \times [1 - \Phi(\theta_i^{(III)} - \gamma_j^{(III)})], \quad (6)$$

$$\Pr(y_{ij} = B) = \Phi(\theta_i^{(I)} - \gamma_j^{(I)})[1 - \Phi(\theta_i^{(II)} - \gamma_j^{(II)})]\Phi(\theta_i^{(III)} - \gamma_j^{(III)}), \quad (7)$$

$$\Pr(y_{ij} = C) = 1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)}), \quad (8)$$

$$\Pr(y_{ij} = D) = \Phi(\theta_i^{(I)} - \gamma_j^{(I)})\Phi(\theta_i^{(II)} - \gamma_j^{(II)})[1 - \Phi(\theta_i^{(III)} - \gamma_j^{(III)})], \quad (9)$$

and

$$\Pr(y_{ij} = E) = \Phi(\theta_i^{(I)} - \gamma_j^{(I)})\Phi(\theta_i^{(II)} - \gamma_j^{(II)})\Phi(\theta_i^{(III)} - \gamma_j^{(III)}), \quad (10)$$

where $\gamma_j^{(I)}$, $\gamma_j^{(II)}$, and $\gamma_j^{(III)}$ are the respective item parameters for the three subprocesses. The respective person parameters are represented by $\theta_i^{(I)}$, $\theta_i^{(II)}$, and $\theta_i^{(III)}$ and assumed to follow a trivariate normal distribution.

Although in both examples the observed polytomous item is decomposed into binary pseudoitems, I note that, in general, it is not necessary to consider binary subprocesses exclusively. All that is required is that the pseudoitems have a smaller number of categories than the observed item. For example, in one of the reported applications, I present an example in which a five-category item is decomposed into two pseudoitems with two and four categories, respectively.

Covariates

It is informative to test a hypothesized multiple-process structure by including covariates that may have subprocess-specific effects. For example, for the two-process model presented in Equations 3–5, we may hypothesize that a covariate z influences Subprocess I:

$$\Pr(y_{ij} = A) = 1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)} + z_{ij}\beta_j^{(I)}), \quad (11)$$

where $\beta_j^{(I)}$ captures the item-specific influence of covariate z (with values specific to person i and item j) on the selection of Response Category A. In general, the same covariates may be included for each latent process to test their process-specific influences. This approach is of particular interest in an experimental study where a treatment is postulated to modify only one latent subprocess. By including the treatment variable as a predictor for each of the subprocesses, we can test whether the treatment affects the targeted subprocess or also other latent response processes.

Aside from their role in validating a hypothesized response structure, covariates are also important in investigating the presence of differential item functioning (DIF). A DIF analysis addresses the question of whether an item performs differently across subgroups of a population (e.g., different age groups) or across contexts (e.g., the position of an item in different item sequences). For example, using DIF methods (Steinberg, 1994, 2001) showed that the serial position of an item in a questionnaire and the pairing of questions in a vignette can affect the parameters of the graded response models. This experimental approach in detecting DIF can provide important insights about

Table 2
Pseudoitems of the Three-Process Model

Category	Item I	Item II	Item III	Category probabilities
A	1	0	0	$\Phi(\theta_i^{(I)} - \gamma_j^{(I)})\,[1 - \Phi(\theta_i^{(II)} - \gamma_j^{(II)})][1-\Phi(\theta_i^{(III)} - \gamma_j^{(III)})]$
B	1	0	1	$\Phi(\theta_i^{(I)} - \gamma_j^{(I)})\,[1 - \Phi(\theta_i^{(II)} - \gamma_j^{(II)})][\Phi(\theta_i^{(III)} - \gamma_j^{(III)})]$
C	0	—	—	$1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)})$
D	1	1	1	$\Phi(\theta_i^{(I)} - \gamma_j^{(I)})\Phi(\theta_i^{(II)} - \gamma_j^{(II)})[1 - \Phi(\theta_i^{(III)} - \gamma_j^{(III)})]$
E	1	1	0	$\Phi(\theta_i^{(I)} - \gamma_j^{(I)})\Phi(\theta_i^{(II)} - \gamma_j^{(II)})\Phi(\theta_i^{(III)} - \gamma_j^{(III)})$

Note. Dashes indicate data are missing by design.

both the causes that lead to DIF and the underlying response process. In contrast, the approach presented here is more descriptive because possible sources of DIF are built into the hypothesized response structure. For example, if DIF is a result of respondents differing in their preference for selecting the *neither disagree or agree* category, then this behavior can be diagnosed and easily interpreted using a two-process model as shown in the next section. Moreover, preferences for the middle categories can also be related to covariates to explain this phenomenon. Both features make the multiprocess approach a useful addition to the current toolbox of DIF methods.

Estimation

The presented multiple-process models can be estimated with latent-variable software programs that allow for missing responses. To apply these programs, one needs to convert the observed item responses to pseudoitem responses, as shown in Tables 1 and 2. Thus, each item is coded into two or more pseudoitems depending on the postulated tree structure. These pseudoitems are then modeled using the one-parameter probit or other item response models. In the reported applications, I use Mplus (Muthén & Muthén, 2010) to estimate the pseudoitem parameters and the covariate effects, as well as the covariance matrix of the random person effects. Appendix A provides more technical detail on the likelihood function and model features. Appendix B contains a numerical example illustrating the pseudoitem conversion as well as the corresponding Mplus code.

Applications

Involvement, Direction, and Intensity in Likert Scales

The bipolar structure of Likert items allows for the measurement of positive and negative responses to attitudinal statements because respondents are asked to select one out of several response categories that are semantically balanced around a midpoint or implied midpoint (e.g., a scale with categories *strongly agree* = A, *agree* = B, *neither disagree nor agree* = C, *disagree* = D, *strongly disagree* = E). Because of the ordinal nature of this response scale, the graded-response model (Samejima, 1969, 1997) seems well suited for the analysis of Likert responses. This item response model assumes that the attitudinal statement gives rise to a latent value that is mapped onto one of the response categories based on a threshold process. I briefly review this approach below and contrast it with a multiple-response-process model that captures a respondent's involvement with the attitudinal issue, the direction of the person's attitude, and the intensity with which this attitude is expressed. The multiple-process representation is illustrated with an application assessing the importance of a firm's ethical behavior in the purchase decision of consumers.

Graded-response model. According to the graded-response model, responses are a result of an underlying normally distributed latent variable that is mapped onto the discrete response scale. Figure 4 illustrates this mapping. A category is selected when the latent value falls between the upper and lower threshold values that define a response category. The probability that a randomly selected person chooses response category $k (k = 1, \ldots, K)$ can then

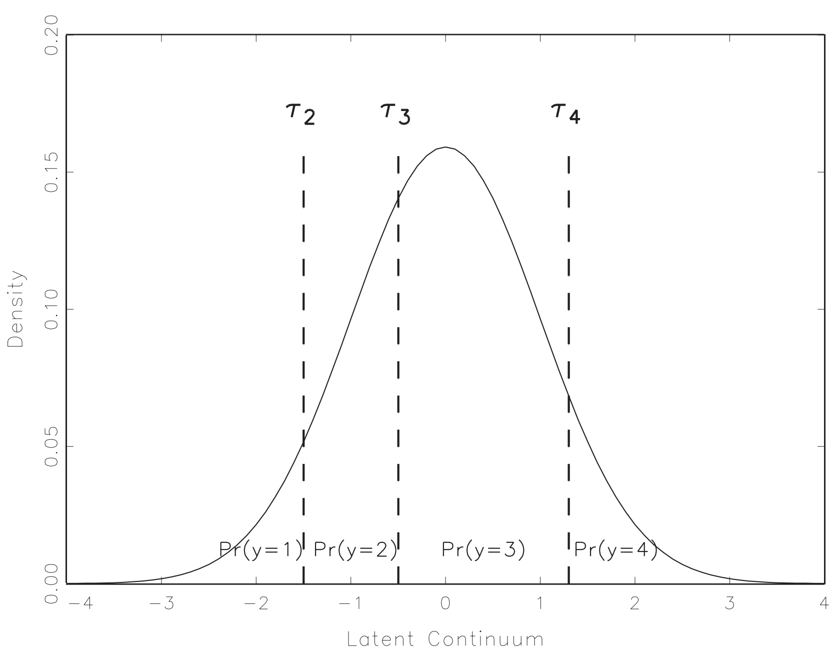

Figure 4. Diagram showing relationship between response probabilities and values of latent variable for Samejima's (1969) model. The response probabilities are defined by the area between two adjacent threshold values τ_k and τ_{k+1} for $k = 1, \ldots, 4$. For example, the probability of observing Response Category 2 is given by the area between τ_2 and τ_3 under the normal density function. The threshold values τ_1 and τ_5 are $-\infty$ and ∞, respectively.

be written as a difference between two normal cumulative distribution functions:

$$\Pr(y_{ij} = k) = \Phi(\tau_{k+1} + \theta_i - \gamma_j) - \Phi(\tau_k + \theta_i - \gamma_j), \quad (12)$$

where τ_k is the threshold value corresponding to the upper boundary of response category k. The lower boundary of category 1 and the upper boundary of category K are minus and plus infinity, respectively, and the threshold values $\tau_2, \tau_3, \ldots, \tau_K$ represent the remaining $K - 1$ boundaries on the latent continuum. McCullagh and Nelder (1989, pp. 151–155) provided more detail on this ordinal response model and related approaches for analyzing ordinal responses.

The next two models allow for a behaviorally more descriptive representation of the response process to Likert items. First, I present a two-process model that allows for items to differ in the degree to which they trigger a respondent to select the midpoint of the scale. Subsequently, I present a three-process model that assumes that the two decisions about which side to take on an issue and how strongly one's opinion is held may be separable and can be influenced by different factors.

A two-process graded-response model. The middle response category, *neither agree nor disagree*, may prompt respondents to think about whether they care or feel strongly about the attitudinal issue (Schwarz & Sudman, 1996). If respondents decide that their opinion is not well developed, they are likely to select the middle category. However, if respondents feel that they have a clear opinion about the issue, they may select one of the remaining response categories according to the graded-response model described above. This two-stage process can be captured by decomposing the observed five-category response into a binary pseudoitem (which codes whether a respondent selects the middle category) and an ordinal four-category pseudoitem. By excluding the indifference category, the latter pseudoitem captures the direction and intensity of the attitudinal self-report. Thus, the probability of selecting the middle category (denoted by C) may be written as

$$\Pr(y_{ij} = C) = \Phi(\theta_i^{(I)} - \gamma_j^{(I)}), \quad (13)$$

and for the remaining four response categories, we obtain

$$\Pr(y_{ij} = A) = [1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)})]\Phi(\tau_B + \theta_i^{(II)} - \gamma_j^{(II)}), \quad (14)$$

$$\Pr(y_{ij} = B) = [1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)})][\Phi(\tau_D + \theta_i^{(II)} - \gamma_j^{(II)}) - \Phi(\tau_B + \theta_i^{(II)} - \gamma_j^{(II)})], \quad (15)$$

$$\Pr(y_{ij} = D) = [1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)})][\Phi(\tau_E + \theta_i^{(II)} - \gamma_j^{(II)}) - \Phi(\tau_D + \theta_i^{(II)} - \gamma_j^{(II)})], \quad (16)$$

and

$$\Pr(y_{ij} = E) = [1 - \Phi(\theta_i^{(I)} - \gamma_j^{(I)})][1 - \Phi(\tau_E + \theta_i^{(II)} - \gamma_j^{(II)})], \quad (17)$$

where $\gamma_j^{(I)}$ and $\gamma_j^{(II)}$ represent the item parameters for the two subprocesses. The respective person parameters are given by $\theta_i^{(I)}$ and $\theta_i^{(II)}$ and assumed to follow a bivariate normal distribution.

A three-process response model. Past work on response style (Baumgartner & Steenkamp, 2001; Cronbach, 1950; Rorer, 1965; Paulhus, 1991) showed that a respondent's use of a rating scale can be determined by variables that are unrelated to the attitude being measured. For example, some respondents may avoid using extreme response categories even if they feel strongly about a topic. Other respondents may tend to select extreme response categories. Predictors of these behaviors include ambivalence avoidance, rigidity, and certainty (De Jong, Steenkamp, Fox, & Baumgartner, 2008). It is important to note that response style effects have been shown to induce spurious correlations among otherwise unrelated constructs (Baumgartner & Steenkamp, 2001; Chun, Campbell, & Yoo, 1974; Hui & Triandis, 1985). To overcome this problem, data from Likert scales are sometimes reduced to the binary level by combining all *agree* and *disagree* responses into two categories: *accept* and *reject*. However, collapsing weak and strong responses is not desirable because potentially valuable information about the intensity of attitudes is lost. I propose an alternative approach by modeling the intensity decision separately from the direction decision. If systematic scale usage biases are present in the data, they can be detected using this approach—and possibly controlled for.

The proposed three-process model captures separate decisions involving the direction and the intensity expression of an attitude. Thus, this model assumes that after individuals take sides on an attitudinal issue (the direction), they next need to decide how strongly their opinion is held. Because decisions on both the direction and intensity of an attitude can be influenced by different factors, it is desirable to model them separately. The response probabilities under this model are given by Equations 6–10.

Do consumers care about business ethics? Investigating whether consumers take (un)ethical activities by businesses into account in their purchase decisions, Creyer and Ross (1997) developed a questionnaire to measure the importance of the ethicality of a firm's behavior. Roux (2006) followed up on this work by administering the questionnaire to a large online sample of Canadian French-speaking consumers. Because her study was conducted in French, the items were translated twice, from English to French and from French to English, to validate the interpretation of the French items. To facilitate reanalysis of these data, I present the frequency distribution of two items from the questionnaire in Table 3. The online supplemental materials contain a complete listing of the Mplus output for this table.

The two selected items (with their respective labels in parentheses) are

- Whether a firm is ethical is not important to me in making my decision what to buy. (ethical)
- Whether a firm is unethical is not important to me in making my decision what to buy. (unethical)

The means of two covariates, age and education, are also included in Table 3. The latter covariate was measured on a 7-point scale ranging from elementary school to doctorate degree.

These items were selected for two reasons. First, they are similar in the sense that they assess whether people value ethical firms. However, they also differ in one respect. In the first question, respondents are asked to think about positive ethical behavior of a firm, whereas in the second question, respondents are asked to consider unethical behavior of a firm. However, this difference is subtle, and one may conjecture that overall both questions may

Table 3
Observed and Predicted Frequencies of Study on Ethical Values

Items					Predictions		
Ethical	Unethical	Age	Education	Observed	1 process	2 process	3 process
SD	SD	46.8	4.0	247	202.5	228.0	244.8
SD	D	42.4	3.4	39	86.3	75.4	41.8
SD	NDA	56.3	3.3	18	35.7	24.5	17.2
SD	A	43.2	3.5	6	2.8	8.7	3.7
SD	SA	50.3	3.2	14	0.5	1.3	16.1
D	SD	47.6	3.8	52	102.3	92.5	51.5
D	D	41.1	3.8	188	117.7	132.8	185.4
D	NDA	43.7	3.2	55	91.0	48.3	65.0
D	A	47.8	4.0	21	17.2	40.2	19.7
D	SA	59.3	3.7	6	4.4	11.9	4.5
NDA	SD	43.9	3.1	18	41.2	22.7	16.1
NDA	D	39.7	3.4	58	91.6	42.8	56.3
NDA	NDA	43.5	3.2	259	154.4	256.2	259.4
NDA	A	43.1	3.5	33	47.4	24.9	36.2
NDA	SA	48.0	3.7	7	19.1	21.2	8.3
A	SD	57.9	3.7	11	5.4	14.8	7.0
A	D	47.9	3.5	32	26.2	56.5	30.4
A	NDA	43.4	3.2	58	69.1	35.1	48.4
A	A	42.0	3.3	63	34.6	35.3	67.4
A	SA	47.0	2.7	7	27.9	20.7	13.1
SA	SD	53.8	3.2	18	0.8	1.8	25.0
SA	D	40.5	3.3	6	5.8	15.9	5.6
SA	NDA	37.4	3.2	9	28.3	27.0	10.2
SA	A	44.7	3.5	11	29.6	20.4	10.7
SA	SA	44.6	3.0	51	45.5	28.0	43.3
LR (*df*)					471.6 (14)	296.5 (11)	16.0 (12)

Note. SD = *strongly disagree*; D = *disagree*; NDA = *neither agree or disagree*; A = *agree*; SA = *strongly agree*; LR = log-likelihood ratio. Age is in years; education is measured on a 7-point scale ranging from 1 = *elementary school* to 7 = *doctorate*.

yield similar answers. Second, in both questions, the importance of ethical behavior is negated, which complicates the thought process (Tourangeau, Rips, & Rasinski, 2000) and may trigger a respondent to select the midpoint of the scale. Some initial support for these observations is provided by a visual inspection of Table 3. Approximately 20% (259) of the 1,287 respondents selected the same response category for both items. The modal response for both items is the midpoint of the scale, indicating that a substantial number of respondents choose to neither disagree nor agree with the two items.

The last three columns in Table 3 contain the predicted frequencies of three models fitted to the data. The last row contains the corresponding log-likelihood ratio (LR) statistics, obtained by comparing observed and predicted frequencies for each model. The first (single-process) model is the graded-response model, which requires the estimation of 10 parameters (one item location, three threshold parameters, as well as the variance of the person parameter for each item). This model yields a poor fit of the data. Specifically, this model does not capture the strong agreement between both items and predicts a less frequent selection of the scale midpoint than is observed in the data. The second (two-process) model allows for the possibility that respondents may first decide whether they feel strongly about the issue. This model requires the estimation of two additional parameters for each item, capturing the degree to which an item triggers indifference. The fit improvement is considerable compared with the graded-response model, but the absolute model fit is still poor, with $LR = 296.5$ and 11 degrees of freedom. As expected, the two-process model does well at predicting the selection of the midpoint for both items. However, it does not improve substantially on the prediction of other cells compared with the graded response model. In contrast, the third (three-process) model performs satisfactorily for the data ($LR = 16$ with 12 degrees of freedom). It captures both the agreement as well as the disagreement between the two ratings. From these results, we can conclude that respondents differ in their degree of involvement, in their attitudinal direction, and in the intensity expression of their attitude. All three subprocesses matter in obtaining a satisfactory description of the data.

Table 4 displays the item parameter estimates of the three-process model for both items and the random effects (see also the online supplemental materials). The two items appear to have similar estimates for the first and third process but differ in their estimates for the second process, with $\gamma_1^{(II)} = -.93$ and $\gamma_2^{(II)} = -1.17$. A simplified model with $\gamma_1^{(I)} = \gamma_2^{(I)}$ and $\gamma_1^{(III)} = \gamma_2^{(III)}$ yields a difference LR test of 3.8 on 2 degrees of freedom. Thus, the items differ only in the extent to which they trigger a directional response. The second item focusing on unethical behavior triggers a more negative attitudinal response than does the first item, which emphasizes the ethical behavior of a firm. However, there are no apparent differences in scale usage and in the extent to which the items elicit a neutral response.

Covariates. To better understand possible sources of individual differences in self-reported ethical consumption behavior, I investigated whether the respondents' educational degree and age

Table 4
Item and Random-Effect Estimates of the Three-Process Model

Effects	Estimates	*SE*
$\gamma_1^{(I)}$	1.10	0.09
$\gamma_2^{(I)}$	1.00	0.09
$\gamma_1^{(II)}$	−0.93	0.15
$\gamma_2^{(II)}$	−1.17	0.16
$\gamma_1^{(III)}$	1.05	0.15
$\gamma_2^{(III)}$	0.93	0.15
$\sigma^{2(I)}$	3.03	0.43
$\sigma^{2(II)}$	5.63	1.11
$\sigma^{2(III)}$	7.40	1.31
$\sigma^{2(I, II)}$	−1.95	0.42
$\sigma^{2(I, III)}$	−3.40	0.47
$\sigma^{2(II, III)}$	2.23	0.48

could account for variability in the three subprocesses identified in the previous analyses. The effects of age are difficult to predict because of two opposing mechanisms. On the one hand, one may conjecture that younger people are more aware of (un)ethical behaviors of firms and are thus more concerned about this issue in their purchase decisions. On the other hand, younger people typically are in a more fragile financial position, which makes extra expenditures for ethically produced products more burdensome. One strategy to cope with this dilemma is to discount unethical firm behaviors. Both mechanisms—greater awareness and greater discounting—can cancel each other, with the result that age may not have a strong predictive effect.

In contrast to age, education has been shown to be predictive of ethical consumption (Starr, 2009). One potential explanation for this finding is that education promotes both thinking about and acting in favor of the public good (Nie, Junn, & Stehlik-Berry, 1996). In my analysis of the effects of age and education, I included both mean-centered covariates at each subprocess level and thus estimated six parameters. This model led to a fit improvement over the model without covariates, with a deviance difference of $\Delta G^2 = 71.5$ and 6 degrees of freedom.

Table 5 shows the regression effects and their standard errors for each of the three subprocesses. Age accounts significantly for variation in the intensity subprocess. Older respondents tend to use the *strongly* response category more often than younger respondents do. However, no significant age effect is observed for the indifference or the direction decision. Thus, age seems to be related predominantly to scale usage—older respondents are less likely to report being unsure—and not to the attitudinal position of the respondents on ethical consumption. The scale-usage effect of age is consistent with a recent study by Soubelet and Salthouse (2011), who found an age-related increase in social desirability. Thus, an interesting follow-up hypothesis is that older respondents tend to be more extreme in their responses on business ethics because they are more prone to give socially desirable responses.

In contrast to age, education appears to have a strong effect on the indifference and direction subprocesses. Participants with higher educational degrees express less indifference (i.e., they select the midcategory response less often) and disagree more often with the two statements. Thus education is strongly associated with respondents' attitudes about a firm's ethical behavior.

It is also instructive to inspect the unconditional and conditional random effects estimated under the three-process model. The variances and covariances of $\theta^{(I)}$, $\theta^{(II)}$, and $\theta^{(III)}$ without and with covariates are

$$\hat{\Sigma} = \begin{pmatrix} 3.03 & -1.95 & -3.40 \\ -1.95 & 5.63 & 2.23 \\ -3.40 & 2.23 & 7.40 \end{pmatrix}, \tag{18}$$

and

$$\hat{\Sigma}^{(\text{Age, Education})} = \begin{pmatrix} 2.92 & -1.78 & -3.29 \\ -1.78 & 5.25 & 2.12 \\ -3.29 & 2.12 & 7.18 \end{pmatrix}, \tag{19}$$

respectively. Perhaps not surprising, individual differences in indifference are negatively correlated with individual differences in direction and intensity, which, in turn, are positively correlated. The small reduction in the random-effect estimates when introducing the two covariates age and education suggests that it is worthwhile to explore other measures to explain why individuals differ in their response behavior.

This application demonstrated that the determinants of the association between the two items are more complex than can be captured by the graded response model or the two-process extension. Specifically, three separable sources of individual differences in the answers to the two items were found. Each item elicited decisions about indifference, direction, and intensity that were shown to be associated differentially with education and age covariates. Models based on a smaller number of response processes did not provide a good fit with the data. Although not discussed in detail, the inclusion of the covariates age and education pointed to DIF and significantly improved the relative fit of the one- and two-process models; however, the absolute fit of these models remained poor. Only the three-process model provided a satisfactory representation of the data.

Contextual and Attribute-Based Information in Choices

Consider choosing among the three wine options A, B, and C in Table 6. If one is not a wine expert, it is tempting to choose the middle option because it seems to provide a compromise between the body and complexity attributes of the wines. In contrast, when only wine bottles A and B or B and C are available, Wine B seems less attractive. This phenomenon, referred to as the *compromise effect* (Simonson, 1989; Simonson & Tversky, 1992), is one of the most robust behavioral phenomena in choice research (Kivetz, Netzer, & Srinivasan, 2004). It is important to note that the compromise effect cannot be accounted for by random utility models because these models require that the share of a choice option can never increase when the choice-set size is increased.

Table 5
Regression Effects of the Three-Process Model

	Age		Education	
Process	Estimate	*SE*	Estimate	*SE*
Indifference (I)	.006	.004	.256	.046
Direction (II)	.001	.006	−.394	.077
Intensity (III)	−.026	.007	−.149	.069

Table 6
Three Wine Options

Brand	Body	Complexity	Price
A	8	9	$19
B	7	7	$19
C	6	6	$19

Note. Body was the perception of texture and weight of the wine in the mouth (rated on a scale of 1–10). Complexity was the perception of multiple layers and nuances of bouquet and flavor in the wine (rated on a scale of 1–10).

However, under the compromise paradigm, it can be shown that an alternative's share often increases relative to other existing alternatives when it becomes an intermediate option in a larger choice set and is reduced when it becomes an extreme option in a smaller choice set.

Preferences for the middle option can be made more or less extreme depending on the motivational focus of the decision maker. Specifically, Mourali et al. (2007) showed that participants with a prevention focus find the middle option more attractive than do participants with a promotion focus. These authors conceptualized choice goals in terms of Higgins's (1997) regulatory focus theory, which classifies them into two broad categories: ideals and oughts. Ideals denote aspirations, hopes, and wishes, whereas oughts stand for responsibilities, obligations, and duties. Higgins's theory posits that ideals and oughts entail distinct self-regulatory systems. Regulation in relation to ideals involves a promotion focus, which is a regulatory state concerned with advancement and accomplishment. In contrast, regulation in relation to oughts involves a prevention focus, which is a regulatory state concerned with protection and safety. Mourali et al. (2007) argued that individuals with a prevention focus—who prefer vigilant strategies of making correct rejections and avoiding mistakes—dislike extreme options (options that are attractive on some but unattractive on other attribute dimensions). Thus, a prevention focus favors the safer compromise options, which offer intermediate levels of all attributes and thus minimize the risk of making a mistake.

Clearly, decision makers can choose among the choice options on the basis of contextual information (Is one of the options a compromise?) or on the basis of the attributes that characterize the choice options (Do I like more body or greater complexity?). To capture these two decision processes, I propose using the two-stage model presented in Figure 2. Here the first decision step is to consider contextual information and to pick (or not) the compromise option, B. If the compromise is not picked, then the next step is to choose between Options A and C on the basis of the preferred attribute. This approach provides a parsimonious framework for testing the effect of contextual sensitivities and attribute preferences on choice.

To test the two-stage model, I used both published and unpublished data from Study 1 reported by Mourali et al. (2007). Here 128 participants were asked to choose their preferred option from three different product categories (toothpastes, printers, and restaurants). Half of the participants were exposed to a promotion-mindset manipulation and the other half to a prevention-mindset manipulation. After they made their selections, the participants filled out a questionnaire reporting their promotion and prevention pride (Higgins et al., 2001). In my analysis of the choice data, I used Equation 11 to capture the first decision step:

$$\Pr(y_{ij} = B) = \Phi(\theta_i - \gamma_B^{(I)} + z_i\,\beta_{\text{ExpProm}}), \tag{20}$$

where j represents the choice set and $\gamma_B^{(I)}$ refers to the attractiveness of the compromise option. The probability of selecting the compromise option was assumed to be the same for the three product categories because it was equally salient in each of them. The mindset manipulation is represented by the dummy variable z_i. This variable takes on the value 0 for the prevention condition and the value 1 for the promotion condition.

Person i's selection of Option A from choice set j at the second step is modeled by the difference between the utilities estimated for Options A and C, denoted by $\gamma_{A(j)}^{(II)}$ and $\gamma_{C(j)}^{(\mathrm{II})}$, respectively.

$$\Pr(y_{ij} = A) = [1 - \Phi(\theta_i - \gamma_B^{(I)} + z_i\,\beta_{\text{ExpProm}})]\Phi(\gamma_{A(j)}^{(II)} - \gamma_{C(j)}^{(II)}). \tag{21}$$

In the data analysis, we set $\gamma_{C(j)}^{(II)} = 0$ because only the difference between the utilities is estimable. Because each option was presented only once, no information is available about individual preference differences that are attribute based. However, because each choice set has a compromise option, we can assess how participants differ in the propensity to select this option, which is captured by the individual difference parameter θ_i. We assume that θ_i follows a normal distribution with standard deviation σ.

Detailed residual analyses showed that the two-stage model provides a good fit with the choice data. The fit is illustrated by the close match between the observed and expected frequencies reported in Table 7. It is interesting that the choice frequency of the compromise option, B, does not vary much across choice sets within the promotion or prevention condition. Equations 20 and 21

Table 7
Observed and Expected Choice Frequencies From Three-Option Sets

	Prev 1		Prev 2		Prom 1		Prom 2		Prev 3		Prom 3	
Option	O	E	O	E	O	E	O	E	O	E	O	E
A	7	9.6	15	12.3	16	15.6	21	19.9	18	17.2	19	21.8
B	39	37.9	30	31.3	36	37.9	31	31.3	39	37.9	33	31.3
C	16	14.5	17	18.4	10	8.5	10	10.8	5	6.9	10	8.9

Note. Prev = prevention; Prom = promotion. O and E represent the observed and expected frequencies under the two-process choice model. The frequencies are listed for each of the three product categories (toothpastes, printers, and restaurants), identified by the numbers 1, 2, and 3, respectively.

capture this feature of the data by specifying that the effect of the compromise option is equal across choice sets. As a result, the expected frequencies for selecting Option B are the same across choice sets within the promotion (31.3) and prevention (37.9) conditions.

The parameter estimates of the two-stage model are reported in Table 8. There are substantial individual differences in the propensity to select the compromise option, with $\hat{\sigma} = 0.50$. The probability of selecting the compromise option for a person with an average propensity is $\Phi(.32) = .63$ in the prevention condition. The difference between the promotion and prevention condition is marginally significant, reducing this probability to about .51, or $\Phi(.32 - .30)$, for participants in the promotion condition. As expected, the probability of preferring Option A over Option C varies from choice set to choice set, with substantial differences among the utility parameters $\hat{\gamma}_{A(j)}$. Figure 5 depicts the choice probabilities of the three options for the first choice set under the promotion condition. This plot illustrates the main features of the model: For low context dependence, preferences of Option A over Option C do not vary. However, with increasing context dependence, Option B becomes increasingly attractive and dominates the other choice outcomes.

As a next step, the participants' promotion and prevention pride scores (Higgins et al., 2001) are included as predictors of the decision to select the compromise option. This model yields a further fit improvement, with a reduction in the deviance of $\Delta G^2 = 27.1$ on 2 degrees of freedom. The obtained parameter estimates are reported in Table 9. Both the promotion and the prevention pride scores are significant. As expected, participants with higher prevention scores are more likely to select the compromise option, whereas participants with higher promotion scores are less likely to select this option. By accounting for individual differences in promotion and prevention focus, we obtain two additional effects. First, the treatment effect reaches significance now because the standard error of this effect is smaller. Second, the unexplained heterogeneity in selecting Option B drops significantly, which demonstrates that the participant-specific preferences for the compromise options across choice sets can be attributed to their regulatory focus.

In sum, this analysis showed that the choice data can be described parsimoniously with a two-stage choice model that allows for random variability in the selection of the compromise option. The two-stage model distinguishes between attribute-based and contextual information as two separate inputs for the choice process. This distinction proved essential in the analysis of the first-choice data. Despite the relatively small sample size, systematic individual differences in contextual sensitivity to the middle option were found, which could be accounted for by the regulatory focus of the participants.

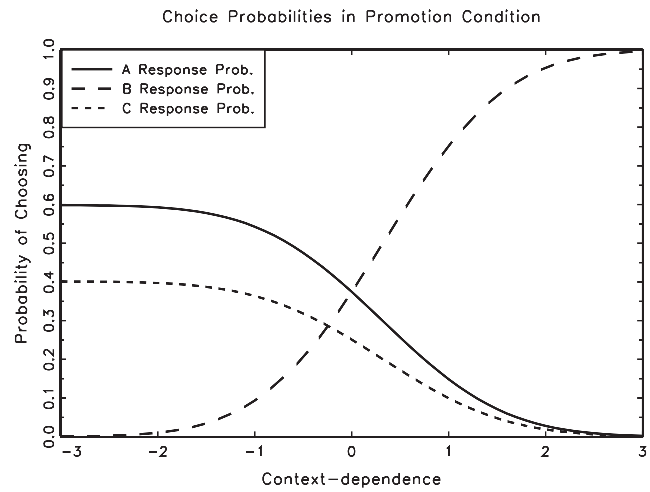

Figure 5. Choice probabilities (Prob.) of the three options A, B, and C estimated under the two-stage model. The probabilities vary as a function of context dependence.

Finally, the two-stage model differs from random-utility models in that it assumes that choices among options are sequential and not simultaneous (Suh & Bolt, 2010). In the reported application, the sequential-process approach proved useful because it allowed the separate effects of contextual and attribute-based information to be distinguished. Both features are difficult to implement in a simultaneous-choice modeling approach. However, it seems premature to favor one framework over the other. More applications are needed to explore sequential and simultaneous influences in a decision process.

Related Models

The presented modeling approach is related to multinomial processing tree (MPT) models (Batchelder, 2009; Batchelder & Riefer, 1999) and diagnostic measurement (DM) models (Rupp et al., 2010). Both classes of models share the notion that variations in categorical outcome measures can be explained by multiple discrete stages or processes. Originally, the development of MPT and DM models was motivated by specific applications in cognitive psychology and educational measure-

Table 8
Item Parameter Estimates of the Two-Process Choice Model

Effects	Estimates	*SE*
γ_{B}	.32	.12
$\gamma_{A(1)}$	.25	.17
$\gamma_{A(2)}$	−.38	.17
$\gamma_{A(3)}$	−.56	.18
$\beta_{ExpProm}$	−.30	.16
σ	.50	.14

Table 9
Parameter Estimates of the Two-Process Choice Model With Covariates

Effects	Estimates	*SE*
γ_{B}	0.32	0.11
$\gamma_{A(1)}$	0.25	0.17
$\gamma_{A(2)}$	−0.38	0.17
$\gamma_{A(3)}$	−0.56	0.18
$\beta_{ExpProm}$	−0.31	0.15
β_{Prom}	−0.06	0.02
β_{Prev}	0.08	0.02
σ	0.30	0.18

ment, respectively. MPT models were used to assess the influence of such cognitive processing capacities as memory storage and memory retrieval. DM models were applied to diagnose discrete skills (or the lack thereof) in solving mathematics items. These different objectives influenced greatly the subsequent refinement and addition of features to both classes of models.

By being tailored to a specific experimental paradigm, the parameters of a MPT model are defined, such that they speak directly to the postulated cognitive architecture in an experimental task. Because the main emphasis in applying MPT models is on testing theories in cognitive psychology, typically MPT models do not include person- or item-specific effects. Instead, MPT analyses focus on counts obtained by summing categorical responses over individuals and items. In contrast, a typical DM application analyzes individual responses scored as correct or incorrect by postulating a set of attributes (in the form of latent classes) that a person must execute to solve an item. For example, to arrive at a correct answer to an algebra problem, basic skills in addition, subtraction, multiplication, or division may be required. Thus, instead of simply scoring how test takers perform on a set of algebra items, DM models allow for the classification of their mastery of a set of discrete skills. Another distinguishing feature between these approaches is that MPT models are sequential in nature, whereas DM models are mostly silent in this regard. However, the importance of this distinction has not been explored. Typically, data collected in MPT and DM applications do not contain information about the temporal nature of the postulated processes. As a result, MPT specifications about the order of cognitive processes are based more on theoretical than on empirical considerations.

Over the years, the scope of MPT and DM approaches has expanded, and they are now applied in a wider range of settings. Recent applications of MPT models include implicit attitude measurements (Sherman et al., 2008), personality assessments (Batchelder, 2009), and the modeling of binary choice data (Batchelder, Hu, & Smith, 2009). Moreover, Klauer (2010) proposed a general Bayesian framework for the estimation of individual differences. DM models were also developed in different directions (for a review, see Rupp et al., 2010), with the result that it has become increasingly difficult to define clearly the methodological boundaries between these two classes of models. One major distinction is still the application goal, with DM models emphasizing diagnostic classification of individuals and MPT models targeting different experimental paradigms in cognitive psychology.

With its focus on response processes to polytomous items, the presented approach is particularly useful for the investigation of judgment and choice processes. The consideration of item- and person-specific effects for each postulated stage, coupled with the inclusion of stage-specific covariates, allows for detailed investigations of well-known but not yet well-understood judgment and choice phenomena. For example, the proposed approach provides a new way to study response-style and response-set effects (Rorer, 1965). Response styles refer to enduring tendencies in answering items that are not specific to the item content. Examples include tendencies to agree with items, to give extreme as opposed to moderate responses, and to give middle or neutral responses (Cronbach, 1950). In contrast to response styles, response sets are related directly to the item content and refer to the motivation of respondents to answer items in a way that facilitates their self-presentation. For example, when measuring compliance with rules and regulations in downloading movies, in online dating settings, or in accurately completing a questionnaire, respondents have been shown to differ not only in their behavior in these different domains but also in their motivation to present themselves in a positive way (Sinha & Mandel, 2008; Toma, Hancock, & Ellison, 2008). There is some evidence to suggest that response-set effects are a result of a stagewise process (Tourangeau et al., 2000), according to which respondents arrive at an initial response on the basis of retrieval processes then subsequently decide whether to edit this response and report a more positive or less revealing answer instead. Using the approach presented here, it is straightforward to implement these stages and test the degree to which positive self-presentation biases affect the observed responses. More generally, modeling how individuals form their responses or choose among options via multiple stages or processes provides a promising methodological approach that may significantly advance the field of judgment and decision making.

Concluding Remarks

In this article, I discussed in detail two applications demonstrating the potential usefulness of modeling multiple response processes in judgment and choice applications via a tree structure. When polytomous items were decomposed into pseudoitems that were fitted, in turn, by item response models, a parsimonious and easily interpretable representation was obtained that provides new insights about how responses are formed to Likert items and how decision makers choose among multiple options depending on their regulatory focus. Moreover, I showed that answers to the pseudoitems can be related to covariates to account for individual differences in each of the multiple response processes. The proposed approach can be implemented readily with current software programs that allow for maximum-likelihood estimation of item response models with missing data.

Although several examples of multiple process models for the analysis of judgment and choice data were presented, this list is by no means exhaustive. Other applications may lead to different tree representations because multiple response processes are likely to depend on the questioning format and the application context. However, as long as the multiple responses can be measured with pseudoitems, the same methods that are discussed here can be used. As illustrated by the two applications, it is straightforward to formulate, estimate, and validate multiple process representations.

More work remains to be done. Perhaps the most critical issue is that an observed response may not always allow for an unambiguous identification of the underlying response process. To some extent, this issue can be addressed by extending item response models for the pseudoitems so they can allow for such process-related features as response times (Klein Entink, Fox, & van der Linden, 2009). For example, when modeling System 1 and System 2 processing in reasoning tasks, one could take into account that immediate responses are faster, on average, than deliberate ones, which would provide additional information about whether an observed response is a result of System 1 or System 2 processing.

A complementary approach is to allow for the possibility that an observed response can be a result of more than one latent response process (Maris, 1995). However, identifiability and estimation are considerably more complex in this case. Still, in view of the potential prevalence of multiple latent responses, these are important topics for future research.

References

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136,* 569–576. doi:10.1037/0096-3445.136.4.569

Batchelder, W. H. (2009). Cognitive psychometrics: Using multinomial processing tree models as measurement tools. In S. E. Embretson (Ed.), *Measuring psychological constructs: Advances in model-based measurement* (pp. 71–94). Washington, DC: American Psychological Association.

Batchelder, W. H., Hu, X., & Smith, J. B. (2009). Multinomial processing tree models for discrete choice. *Zeitschrift für Psychologie/Journal of Psychology, 217,* 149–158.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86. doi:10.3758/BF03210812

Baumgartner, H., & Steenkamp, J.-B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research, 38,* 143–156. doi:10.1509/jmkr.38.2.143.18840

Birnbaum, A. (1968). Some latent trait models and their uses in inferring an examinee's ability. In F. Lord & M. Novick (Eds.), *Statistical theories of mental test scores* (pp. 395–479). Boston, MA: Addison-Wesley.

Böckenholt, U. (2006). Thurstonian-based analyses: Past, present and future utilities. *Psychometrika, 71,* 615–629. doi:10.1007/s11336-006-1598-5

Böckenholt, U. (2012). The cognitive-miser model: Testing for intuitive and deliberate reasoning. *Psychometrika, 77,* 388–399.

Chun, K., Campbell, J., & Yoo, J. (1974). Extreme response style in cross-cultural research: A reminder. *Journal of Cross-Cultural Psychology, 5,* 465–480. doi:10.1177/002202217400500407

Creyer, E. H., & Ross, W. T. J. (1997). The influence of firm behavior on purchase intention: Do consumers really care about business ethics? *Journal of Consumer Marketing, 14,* 421–432. doi:10.1108/07363769710185999

Cronbach, L. (1950). Further evidence on response set and test design. *Educational and Psychological Measurement, 10,* 3–31. doi:10.1177/001316445001000101

De Jong, M. G., Steenkamp, J.-B. E. M., Fox, J. P., & Baumgartner, H. (2008). Using item response theory to measure extreme response style in marketing research: A global investigation. *Journal of Marketing Research, 45,* 104–115. doi:10.1509/jmkr.45.1.104

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59,* 255–278. doi:10.1146/annurev.psych.59.103006.093629

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19,* 25–42. doi:10.1257/089533005775196732

Glas, C. A. W. (2010). Testing fit to IRT models for polytomously scored items. In M. L. Nering & R. Ostini (Eds.), *Handbook of polytomous item response theory models* (pp. 185–210). New York, NY: Routledge.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52,* 1280–1300.

Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology, 31,* 3–23. doi:10.1002/ejsp.27

Hui, H., & Triandis, H. (1985). The instability of response sets. *Public Opinion Quarterly, 49,* 253–260. doi:10.1086/268918

Kivetz, R., Netzer, O., & Srinivasan, V. (2004). Alternative models of capturing the compromise effect. *Journal of Marketing Research, 41,* 237–257. doi:10.1509/jmkr.41.3.237.35990

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75,* 70–98. doi:10.1007/s11336-009-9141-0

Klein Entink, R. H., Fox, J.-P., & van der Linden, W. J. (2009). A multivariate multilevel approach to simultaneous modeling of accuracy and speed on test items. *Psychometrika, 74,* 21–48. doi:10.1007/s11336-008-9075-y

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. *Psychological Science, 8,* 60–64. doi:10.1111/j.1467-9280.1997.tb00545.x

Maris, E. (1995). Psychometric latent response models. *Psychometrika, 60,* 523–547. doi:10.1007/BF02294327

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47,* 149–174. doi:10.1007/BF02296272

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models.* London, England: Chapman & Hall.

Mourali, M., Böckenholt, U., & Laroche, M. (2007). Compromise and attraction effects under prevention and promotion motivations. *Journal of Consumer Research, 34,* 234–247. doi:10.1086/519151

Muthén, B., & Muthén, L. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Nie, N., Junn, J., & Stehlik-Berry, K. (1996). *Education and democratic citizenship in America.* Chicago, IL: University of Chicago Press.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609. doi:10.1037/0022-3514.46.3.598

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. Shaver, & L. Wright (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Rorer, L. G. (1965). The great response-style myth. *Psychological Bulletin, 63,* 129–156. doi:10.1037/h0021888

Roux, C. (2006). *Typologie de consommateurs "conscientisés"* [Typology of "conscientious" consumers] (Unpublished master's thesis). HEC Montréal, Montréal, Quebec, Canada.

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York, NY: Guilford Press.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement, 34*(4, Pt. 2).

Samejima, F. (1997). Graded response model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). New York, NY: Springer.

Schwarz, N., & Sudman, S. (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research.* New York, NY: Jossey–Bass.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115,* 314–335. doi:10.1037/0033-295X.115.2.314

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16,* 158–174. doi: 10.1086/209205

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29,* 281–295. doi:10.2307/3172740

Sinha, R. K., & Mandel, N. (2008). Preventing music piracy: The carrot or the stick? *Journal of Marketing, 72,* 1–15. doi:10.1509/jmkg.72.1.1

Soubelet, A., & Salthouse, T. A. (2011). Influence of social desirability on

age differences in self-reports of mood and personality. *Journal of Personality, 79,* 741–762. doi:10.1111/j.1467-6494.2011.00700.x
Starr, M. A. (2009). The social economics of ethical consumption: Theoretical considerations and empirical evidence. *Journal of Socio-Economics, 38,* 916–925. doi:10.1016/j.socec.2009.07.006
Steenkamp, J.-B. E., De Jong, M., & Baumgartner, H. (2010). Socially desirable response tendencies in survey research. *Journal of Marketing Research, 47,* 199–214. doi:10.1509/jmkr.47.2.199
Steinberg, L. (1994). Context and serial-order effects in personality measurement: Limits on the generality of measuring changes the measure. *Journal of Personality and Social Psychology, 66,* 341–349. doi:10.1037/0022-3514.66.2.341
Steinberg, L. (2001). The consequences of pairing questions: Context effects in personality measurement. *Journal of Personality and Social Psychology, 81,* 332–342. doi:10.1037/0022-3514.81.2.332
Suh, Y., & Bolt, D. M. (2010). Nested logit models for multiple-choice item response data. *Psychometrika, 75,* 454–473. doi:10.1007/s11336-010-9163-7
Toma, C. L., Hancock, J., & Ellison, N. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34,* 1023–1036. doi:10.1177/0146167208318067
Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response.* New York, NY: Cambridge University Press.
Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychological Review, 117,* 1275–1291. doi:10.1037/a0020580
Tutz, F. (1997). Sequential models for ordered responses. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 139–152). New York, NY: Springer.
van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory.* New York, NY: Springer.
Verhelst, N. D., Glass, C. A. W., & de Vries, H. H. (1997). A steps model to analyze partial credit. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 123–138). New York, NY: Springer.

Appendix A

Likelihood Function

As shown in the Multiple-Response-Process Models section, the probability of observing a response category can be written as a product of the pseudoitem probabilities that are associated with the latent branch, leading to the observed outcome. For example, for trichotomous items, the category probabilities are given by Equations 3–5 and for Likert items with five response categories by Equations 6–10.

For an item with K response categories, we can estimate up to $K - 1$ stage-specific person parameters. Thus, the dimensionality of the random effects does not depend on the number of items being considered. In general, for J items and n individuals, the log-likelihood function of the observed responses $\mathbf{y}$ under the standard assumption of conditional independence given the person parameters $\boldsymbol{\theta}_i = (\theta_i^{(I)}, \theta_i^{(II)}, \ldots, \theta^{(K-1)})$ is given by

$$l(\psi|\mathbf{y}) = \sum_{i=1}^{n}\sum_{j=1}^{J}\int\left[\prod_{k=1}^{K}\{Pr(y_{ij} = k \mid \boldsymbol{\theta}_i)\}^{\delta_{kij}}\right]p(\boldsymbol{\theta}_i \mid \eta_{\boldsymbol{\theta}})d\boldsymbol{\theta}_i, \tag{1}$$

where $\delta_{kij} = 1$ if $y_{ij} = k$ and 0 otherwise for $k = 1, \ldots, K$, and η_θ contains the random-effects distribution parameters.

By assuming that each response category has a unique path, the multiple-process models can be estimated easily. As shown in the next section, Mplus or item response theory packages that allow for missing data can be used for this purpose. The fit of multiple-process models can be assessed in the same way as single-process item response theory models. For a small number of items, observed and expected response frequencies can be compared using likelihood ratio tests. Glas (2010) reviewed alternative methods for assessing item and person fit when frequency tables are sparse.

When the number of random effects is small, it is useful to plot the category-specific trace lines. Consider, for example, a scenario in which individuals are asked several sensitive questions about their compliance behavior with certain rules and regulations. Occasionally, these individuals are noncompliant, but instead of giving a truthful ("yes") answer, they refuse to answer the question. A two-stage process is well suited to describe this response behavior. At the first stage, a person gives a "no" response when he or she has been compliant. If the person has been noncompliant, he or she decides to either report the noncompliant behavior or refuse to answer the question at the second stage. Thus, there are two sources of individual differences. The first source refers to the degree of compliance with the domain under study, and the second source refers to the propensity of individuals to answer the question truthfully. Figure A1 provides a graphical representation of these response tendencies in the extreme case when noncompliance is perfectly correlated with a person's propensity to refuse answering the sensitive question. The left panel of Figure A1 shows the probabilistic relationships between giving a "no" response and the noncompliance level of the respondents at the first stage and between refusing to answer the question and the noncompliance level of the respondents at the second stage. The right panel depicts the three category probabilities "no," "yes," and "response refusal" for different noncompliance levels. The trace line of the "yes" response probabilities is single peaked. Thus, noncompliant respondents may initially admit to their behavior but then they may become increasingly more likely to refuse to answer the question. A measurement model under the standard assumption of a monotonic relationship between noncompliance and the probability of a "yes" response would be misspecified and could provide only a biased representation of this process.

(*Appendices continue*)

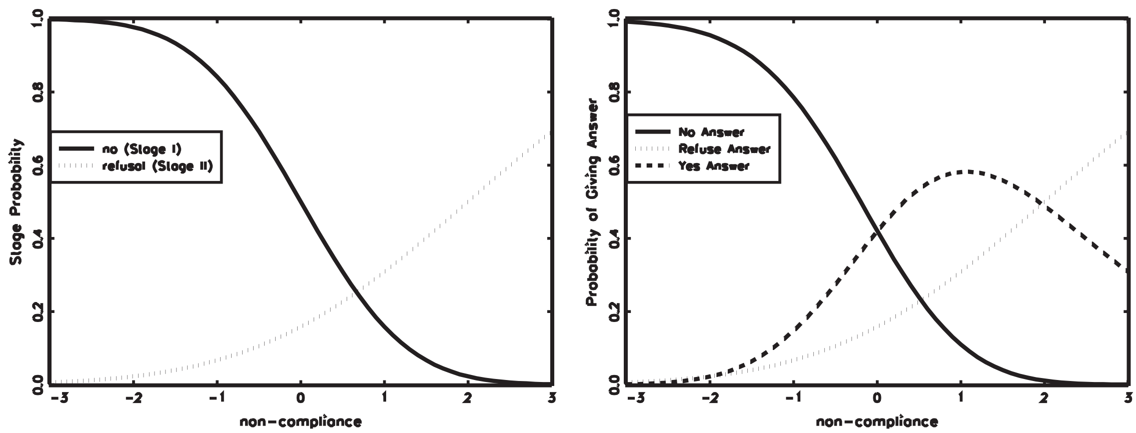

Figure A1. Stage and item-category probabilities of response refusal example. The left panel shows the item trace lines of a "no" response at Stage I and response refusal at Stage II for different noncompliance levels. The right panel depicts the trace lines of the three response categories: "no," yes," and "response refusal." The item trace line of the "yes" response is single peaked, although the trace line of a "no" response is monotonically related to noncompliance.

Appendix B

Mplus Code

To estimate the Multiple-Response-Process model from the discrete judgment or choice data using item response software, one first needs to convert the ratings or choices to pseudoitems that describe the response processes. Examples of this conversion are shown in Tables 1 and 2. For example, for the analysis of the Likert items, each rating is recoded as responses to three binary items that can take on missing values as well. Thus, in the business ethics application, the two Likert items are converted into six binary items that are labeled ×11 and ×12 for the first response process, ×21 and ×22 for the second response process, and ×31 and ×32 for the third response process. Values that are missing by design are coded −9. In addition, the two covariates age and educational degree (abbreviated *edu*) are included in the data set.

Below I provide the Mplus (Muthén & Muthén, 1998–2010) code for the analysis of the business ethics ratings, including covariates. This program estimates the random effects by numerical integration. Because the random effects are specified to follow a multivariate normal distribution, Gauss–Hermite integration with 15 quadrature points is used. The item loadings are constrained to be 1, which allows for the estimation of the full covariance matrix of the random effects. The last line of the code specifies the estimation of the covariate effects on each of the subprocesses defined by the pseudoitems. The online supplemental materials contain the full Mplus output for the data in Table 3.

```
VARIABLE:
  NAMES ARE ×11 ×12 ×21 ×22 ×31 ×32 age edu;
  CATEGORICAL ARE ×11 ×12 ×21 ×22 ×31 ×32;
  USEVARIABLES ARE ×11 ×12 ×21 ×22 ×31 ×32
  age edu;
MISSING ARE ALL (-9);
ANALYSIS:
  ESTIMATOR = MLR; LINK = PROBIT; INTEGRA-
  TION = GAUSSHERMITE(15);
MODEL:
  f1 BY ×11@1 ×12@1; f2 BY ×21@1 ×22@1; f3 BY
  ×31@1 ×32@1;
  f1 ON age edu; f2 ON age edu; f3 ON age edu;
```

Received February 14, 2011
Revision received July 19, 2011
Accepted January 19, 2012 ■