

Article

Personality Traits Leading Respondents to Refuse to Answer a Forced-Choice Personality Item: An Item Response Tree (IRTtree) Model

Martin Storme ^{1,*} , Nils Myszkowski ² , Emeric Kubiak ³  and Simon Baron ³

¹ IESEG School of Management, Université de Lille, CNRS, UMR 9221-LEM-Lille Economie Management, 59000 Lille, France

² Department of Psychology, Pace University, New York, NY 10038, USA

³ AssessFirst, Science and Innovation, 75002 Paris, France; ekubiak@assessfirst.com (E.K.); sbaron@assessfirst.com (S.B.)

* Correspondence: m.storme@ieseg.fr

Abstract: In the present article, we investigate personality traits that may lead a respondent to refuse to answer a forced-choice personality item. For this purpose, we use forced-choice items with an adapted response format. As in a traditional forced-choice item, the respondent is instructed to choose one out of two statements to describe their personality. However, we also offer the respondent the option of refusing to choose. In this case, however, the respondent must report a reason for refusing to choose, indicating either that the two statements describe them equally well, or that neither statement describes them adequately. We use an Item Response Tree (IRTtree) model to simultaneously model refusal to choose and the reason indicated by the respondent. Our findings indicate that respondents who score high on openness are more likely to refuse to choose, and they tend to identify more often with both statements in the forced-choice item. Items containing non-socially desirable statements tend to be skipped more often, with the given reason being that neither proposition describes the respondent well. This tendency is stronger among respondents who score high on agreeableness, that is, a trait that is typically related to social desirability. We discuss the theoretical and practical implications of our findings.

Keywords: personality; forced-choice; item skipping; item response tree



Citation: Storme, M.; Myszkowski, N.; Kubiak, E.; Baron, S. Personality Traits Leading Respondents to Refuse to Answer a Forced-Choice Personality Item: An Item Response Tree (IRTtree) Model. *Psych* **2024**, *6*, 100–110. <https://doi.org/10.3390/psych6010006>

Academic Editor: Alexander Robitzsch

Received: 28 November 2023

Revised: 4 January 2024

Accepted: 8 January 2024

Published: 10 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Forced-choice personality questionnaires are assessment tools that present individuals with sets of statements and, most commonly, ask them to rank the statements according to how well they describe their personality [1,2]. This approach contrasts with the single-statement approach, which usually requires individuals to rate themselves on a scale of how much they agree or disagree with statements that are presented independently from one another [1]. Using forced-choice items is often presented as a way to counter response biases such as the acquiescence and the extremity biases [2]. In personnel selection situations, the forced-choice approach is sometimes preferred to the rating scale approach to assess personality because the former may be less prone to faking or social desirability biases, as individuals are thought not to be given the opportunity to present themselves in a more favorable light by choosing responses that align with their desired image [3]. There is evidence that questionnaires using forced-choice items are indeed less prone to faking than questionnaires using traditional single-statement items [4,5]. Research also suggests that forced-choice items exhibit strong equivalence with traditional single-statement items, demonstrating similar reliabilities, validities, and minimal differential impact on respondents' emotional and cognitive reactions [6].

With forced-choice items, just like with single-statement items, it is theoretically possible to leave the respondent the choice of not choosing a response, that is, skipping

an item. Giving the choice to respondents to provide or not provide a response to an item allows respondents not to answer questions they are uncomfortable with or do not have a response for. This aligns with the general ethical principles of respecting individual autonomy in personality assessments [7]. Beyond the ethical issue, forcing respondents to answer questions they do not have a response to could result in misleading or inaccurate responses. It is generally important to strike a balance between collecting as much information as possible while also ensuring that the responses obtained are meaningful and representative of the respondent's beliefs and attitudes.

While it may seem important to leave the respondent the choice of not choosing a response option, not choosing can adversely affect the psychometric properties of the assessment. These non responses can be viewed as missing data, resulting in reduced reliability of the measurement and larger standard errors of measurement when using Item Response Theory (IRT) factor scoring. What is also problematic is that the decision of refusing to choose an answer in a forced-choice item could be linked to characteristics of items or specific personality traits of the respondents—possibly the very ones that are supposed to be assessed by the questionnaire. In order to better understand what leads some individuals to skip a forced-choice item, it is important to study the personality factors contributing to this phenomenon in interaction with item characteristics.

1.1. Why Would Respondents Refuse to Choose?

To understand the mechanisms that could lead some respondents to refuse to give a response to a forced-choice item, it is useful to first describe the process of responding to a forced-choice item. Ref. [1] suggested using Thurstone's law of comparative judgment [8] to model the response process to forced-choice items. Specifically, the respondent first assigns a utility value to each statement composing an item—the utility function being a linear function of the latent trait [1,2]. The utilities of different statements are then compared by the respondent who ranks the statements according to their utility [1,2]. In the case of items presenting two statements, for example, the chosen statement will be the one with the highest utility.

We propose two main psychological mechanisms to explain individual differences in the propensity to refuse to respond to a forced-choice item. The first mechanism relates to some individuals having more complex self-schemas than other individuals. The complexity of self-schemas is sometimes referred to as self-complexity, that is, the number of self-aspects that compose one's identity [9]. Individuals with higher levels of self-complexity have a self concept that is composed of multiple, distinct facets or identities, rather than being overly simplified or one-dimensional. Following the definition of the construct, it should be more difficult for someone with more complex self-schemas to rank personality descriptors based on their utility than for someone with simpler self-schemas. This is because individuals with complex self-schemas recognize that several personality descriptors may accurately represent different facets of their multifaceted self. The utility values of the statements composing an item may appear to them as indistinguishable. Consequently, they may subjectively experience that several statements describe them equally well, making it difficult for them to rank the statements. Refusing to respond to forced-choice items could thus be thought as a consequence of having complex self-schemas.

If refusing to choose a response can be thought of as a consequence of having complex self schemas, one might expect that the personality traits that typically correlate with complex thinking would also correlate with the tendency to skip forced-choice items. Openness is the general personality trait that is most consistently associated with indicators of complex thinking such as need for cognition [10], ambiguity tolerance [11], and integrative complexity [12]. Individuals who score high on openness can thus be expected to more often experience a difficulty to rank statements in a forced-choice item because of the complexity of their self perception and their tendency to recognize many statements as descriptive of who they are. Consequently, we hypothesized that openness is positively associated with the propensity to skip forced-choice items.

There is a second mechanism related to social desirability that could explain why some respondents refuse to give a response to a forced-choice item. Social desirability is defined as the tendency to both overstate positive characteristics and understate negative ones. Research shows, however, that there is an asymmetry between these two tendencies [13]. Specifically, it appears that social desirability is more about understating negative characteristics than overstating positive ones [13]. In the context of an assessment, it should thus be more important for a respondent not to present themselves in a negative light than to present themselves in a positive light. When the statements composing a forced-choice item are socially undesirable, respondents are forced to describe themselves in a socially undesirable way, which should be particularly threatening with regard to social desirability. For items using single-statements with a rating scale, the respondent can always indicate that the item does not describe them; however, with forced-choice items, there is no escape. To avoid having to describe themselves using undesirable statements, respondents could be tempted to skip the item. It should be especially the case for respondents who have a strong need for social desirability.

We can expect that personality traits that are associated with a strong propensity for social desirability will be linked to skipping forced-choice items containing socially undesirable statements. The literature on the links between HEXACO and social desirability suggests that all traits are related to social desirability, except for openness [14]. Consequently, it can be hypothesized that, except for openness, all of the HEXACO traits (honesty/humility, emotional stability, extroversion, agreeableness, and conscientiousness) will be more strongly associated with item skipping among items containing socially undesirable statements than among items containing socially desirable statements. For individuals scoring high on these personality traits, refusing to give a response to items that are not socially desirable would be a way to avoid having to present themselves in a less favorable light. In sum, we hypothesize that traits associated with social desirability will interact with item desirability in predicting the propensity to skip forced-choice items.

1.2. Study Overview

To test our hypotheses, we used a series of sixty forced-choice items with an adapted response format. Each item contains two personality descriptors drawn from the HEXACO model. Within the same item, the descriptors are either both positive (e.g., honest, confident, extrovert, friendly, organized, curious), or both negative (e.g., dishonest, nervous, introvert, challenging, careless, cautious). In all, the questionnaire contains thirty positive and thirty negative items. For each item, the respondent is asked to choose the descriptor that best describes their personality. The respondent may also choose not to select one of the two descriptors. In this case, the respondent is asked to indicate whether the two descriptors describe them equally well, or whether neither proposition describes their personality.

In the present study, we are interested in modeling (1) the refusal to choose and (2) the reason given for the refusal to choose. In the case where the respondent decides to choose one of the two descriptors, we do not model this choice. The reason for this is that, to our knowledge, there is currently no model that can take into account both the process of responding to a forced-choice item—which requires the use of a Thurstonian IRT model [15]—and the process of refusing to respond—which requires the use of an Item Response Tree (IRTtree) model [16]. Validating such a model is beyond the scope of this paper. In the current paper, we use an IRT model [16] to simultaneously model the refusal to choose and the reason given by the respondent for skipping. IRTtree models combine elements of Item Response Theory (IRT) and decision tree modeling to provide a more flexible and informative approach when modeling complex response processes [16]. Specifically, the response process is decomposed into more basic processes and modeled using nodes—sometimes referred to as pseudo-items [17].

For the purpose of our study, we propose to decompose the process of skipping items in our questionnaire into two more basic processes. The first one is that the respondent can decide to choose one of the two descriptors or not (Node 1). If the respondent decides not

to give an answer, the respondent can then choose to indicate that none of the descriptors describes them or that both descriptors describe them (Node 2). Note that we do not make assumptions about the order in which these two processes happen and, even if we wanted to, an IRTree model gives no information on the order of processes [17]. To investigate the influence of respondents' personality traits and item positivity on the propensity to skip items for one reason or another, we include personality and item covariates and their interactions with the IRTree model. A conceptual and simplified version of the model is presented in Figure 1.

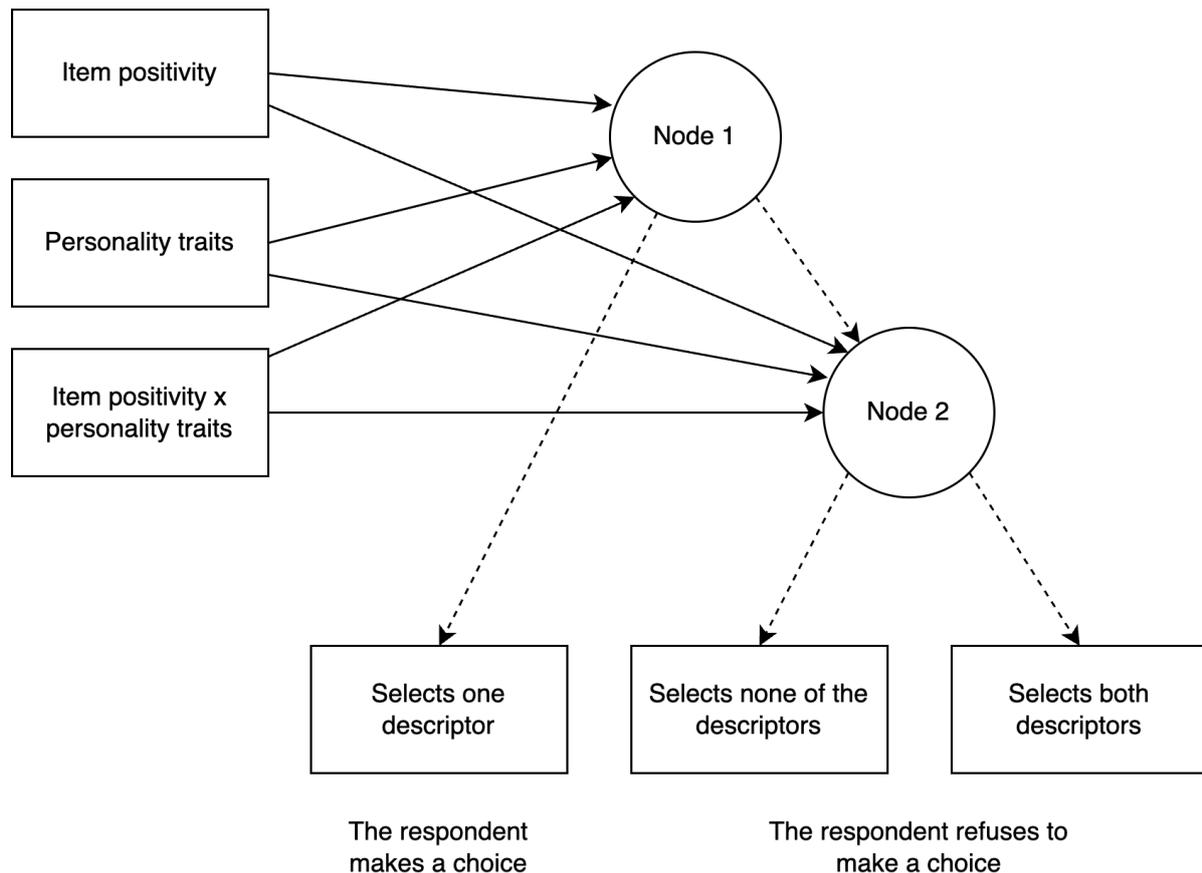


Figure 1. Conceptual and simplified version of the IRTree model used in the current study.

2. Methods

2.1. Participants and Procedure

The sample for this study consisted of 497 users of an online recruitment platform ($M_{age} = 42.84$, $SD_{age} = 11.24$, 325 women, 164 men, and 8 participants who identified as non-binary or prefer not provide information about their gender). We invited users who had filled in a forced-choice personality questionnaire as part of building their online personal profile to fill in an additional personality questionnaire using rating scales for a research project via an online survey. No respondent was excluded from the analyses. All respondents provided informed consent to participate in the study.

2.2. Measures

We administered a personality questionnaire using forced-choice items to measure participants' skipping. The questionnaire was developed using the assessment platform and assesses personality dimensions according to a model based on HEXACO—honesty/humility, emotionality (i.e., the opposite of emotional stability), extroversion, agreeableness, conscientiousness, and openness. It consists of sixty pairs of statements (half of them are positive) for which users have to indicate which one describes them best. To classify items as positive

or negative, we relied on a recent large language model that has been fine-tuned to provide social desirability ratings of questionnaire items, including personality items [18]. In the validation study, the authors reported a correlation of 0.80 between human ratings and machine ratings on a validation sample of items, making the machine a cost effective alternative to human raters [18]. In our study, items that received a positive social desirability score from the machine were coded as positive, whereas items that received a negative social desirability score were coded as negative. Users also have the possibility not to make a choice by selecting one of the following two non-focal response options: “none of the options describes me well” or “both options describe me equally well”.

For our analyses, the personality traits and facets of respondents were not assessed using the questionnaire described above. Instead, we used the Big Five Inventory 2 [19,20], which is a more extensively researched instrument, and, due to its Likert scale response format, facilitates the investigation of between-individual differences compared to a forced-choice questionnaire. This questionnaire consists of 60 statements—we added 12 items to measure honesty/humility (72 items in total)—for which respondents report their level of agreement. The questionnaire assesses honesty/humility (facets: greed avoidance, modesty, and sincerity), emotionality (facets: anxiety, depression, and emotional volatility), extroversion (facets: sociability, assertiveness, and energy level), agreeableness (facets: compassion, respectfulness, and trust), conscientiousness (facets: organization, productiveness, and responsibility), and openness (facets: aesthetic sensitivity, intellectual curiosity, and creative imagination). The scale-score reliability was satisfactory for all six general personality traits with Cronbach’s α ranging between 0.73 and 0.88.

2.3. Statistical Analyses

In the statistical analyses, we used Item Response Tree (IRTree) models to account simultaneously for the refusal to choose (Node 1) and the reason given for the refusal to choose (Node 2) [16]. Because we are interested in simultaneously including person and item covariates as predictors in the model, we used the multidimensional random item model for linear response trees [21]. With this model, the variance and covariance of the latent variables—that is, the tendency to skip and the tendency to choose one reason over another—are estimated for both items and persons [21].

We estimated a series of models with and without person and item covariates and compared their fit using the Akaike Information Criterion (AIC) [22]. Since our planned comparisons involved nested models, we also used likelihood ratio tests to compare models. Model 1 consisted of a baseline model that did not include any person or item covariates. In Model 2, we added the node, the positivity of the item (effect coded: -0.5 for negative and $+0.5$ for positive) and the six HEXACO standardized scores as fixed effects. Finally, in Model 3, we added all two-way interactions between the node and item positivity (effect coded: -0.5 for negative and $+0.5$ for positive), all two-way interactions between the node and each one of the six HEXACO standardized scores, and all three-way interactions between the node, item positivity, and the six HEXACO standardized scores. For each model, we report the marginal and conditional coefficients of determination for generalized mixed-effect models (R_{GLMM}^2) estimated with the R package MuMIn [23] based on [24]. The marginal R_{GLMM}^2 represents the variance explained by the fixed effects, while the conditional R_{GLMM}^2 is interpreted as the variance explained by the entire model, including both fixed and random effects [23,24].

The dataset was prepared using the *irtrees* package [21], a package that provides a series of functions to reorganize datasets and facilitate the estimation of tree-based item response models with the R package *lme4* [25]—a statistical modeling package for estimating generalized linear mixed models (GLMMs). To estimate model parameters, we relied on *lme4* along with its popular extension *lmerTest* [26]. GLMMs and *lme4* have been discussed as particularly useful to estimate IRT models [27], particularly IRTree models [21].

3. Results

As expected, Model 1 was outperformed by Model 2, $\chi^2(8) = 40.47$, $p < 0.001$, $AIC_1 = 29,076.47$, $AIC_2 = 29,052.00$, which was, in turn, outperformed by Model 3, $\chi^2(19) = 580.41$, $p < 0.001$, $AIC_2 = 29,052.00$, $AIC_3 = 28,509.59$. This indicates that there are significant effects of the person and item covariates and their interactions on the propensity to refuse to choose a response option and to give one of two reasons as a justification. Consequently, we report the estimates of Model 3 in all subsequent analyses (see Table 1).

Table 1. Estimates of the fixed effects of all tested models.

Covariate	Model 1	Model 2	Model 3
Intercept	−1.563 ***	−1.786 ***	−1.772 ***
Node		1.069 *	1.134 **
Item positivity		0.509 **	−1.207 ***
Humility		0.077	0.271 *
Negative Emotionality		0.075	0.076
Extraversion		−0.050	0.061
Agreeableness		0.090	0.181
Conscientiousness		0.103	0.176
Openness		0.297 ***	0.279 **
Node × Item positivity			6.330 ***
Node × Humility			−0.298 *
Node × Negative Emotionality			−0.002
Node × Extraversion			−0.159
Node × Agreeableness			−0.140
Node × Conscientiousness			−0.115
Node × Openness			0.032
Item positivity × Humility			0.017
Item positivity × Negative Emotionality			0.274 ***
Item positivity × Extraversion			−0.073
Item positivity × Agreeableness			−0.129 **
Item positivity × Conscientiousness			−0.011
Item positivity × Openness			−0.039
Node × Item positivity × Humility			0.001
Node × Item positivity × Negative Emotionality			−1.221 ***
Node × Item positivity × Extraversion			0.479 ***
Node × Item positivity × Agreeableness			0.906 ***
Node × Item positivity × Conscientiousness			0.356 **
Node × Item positivity × Openness			0.595 ***
Akaike Information Criterion	29,076.47	29,052.00	28,509.59
Marginal R^2_{GLMM}	0.00	0.03	0.19
Conditional R^2_{GLMM}	0.61	0.60	0.61

The reference level for the Node variable is Node 1. Item positivity is effect-coded. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Our findings suggest that the overall probability to skip a forced-choice item is 0.15. Given that the item is skipped, the reason that is most often reported is that none of the statements adequately describes the respondent, with an overall probability of 0.65. In line with our hypotheses, we found that the level of openness of the respondent was positively associated with the tendency to refuse to choose one of the statements, $B = 0.28$, $z = 3.08$, $p < 0.01$. Openness also appears to be associated with the reason given for the refusal to choose, with respondents scoring high on openness being more likely to refuse to choose because the two statements describe them well rather than respondents scoring low on openness, $B = 0.32$, $z = 4.34$, $p < 0.001$ (see Figure 2).

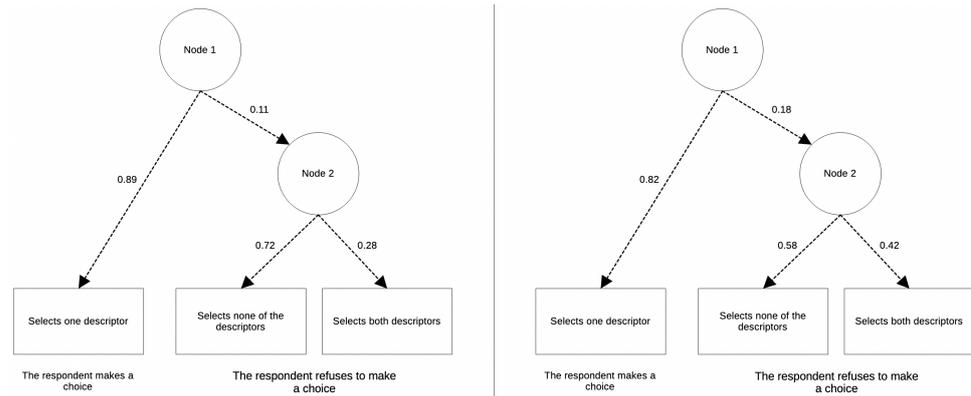


Figure 2. Item skipping and given reason at different levels of openness. (**Left panel**): The respondent scores low on openness (-1 SD). (**Right panel**): The respondent scores high on openness ($+1$ SD)

In line with our hypotheses, items consisting of positive statements were skipped less often than items consisting of negative statements, $B = -1.21$, $z = -4.61$, $p < 0.001$. The tendency to skip negative items more often than positive items appears to be stronger among respondents who score high on agreeableness, $B = -0.13$, $z = -3.15$, $p < 0.01$, and emotional stability, $B = -0.27$, $z = -7.11$, $p < 0.001$, and marginally stronger among respondents who score high on extroversion, $B = -0.07$, $z = -1.86$, $p = 0.06$. As expected, openness was not associated with the tendency to skip negative items, $B = -0.04$, $z = -1.15$, $p = 0.25$. However, contrary to our hypotheses, the tendency to skip negative items appears to be unrelated to honesty/humility and conscientiousness.

Interestingly, the reason given to skip negative items tends to be different from the reason given to skip positive items, $B = 5.12$, $z = 8.90$, $p < 0.001$. For positive items, the probability that respondents indicate that both statements describe them equally well when they refuse to choose is 0.87, whereas for negative items, it is only 0.04. The difference in the reasons provided to justify item skipping is also linked to the respondent's level of agreeableness, $B = 0.78$, $z = 6.70$, $p < 0.001$ (see Figures 3 and 4).

In response to a reviewer's suggestion, we conducted a supplementary analysis wherein we collapsed the two reasons for refusing to answer a forced-choice item, aligning with the structure of the analysis at Node 1 in Model 3. The results of this simplified analysis are in line with Model 3. Indeed, the simpler analysis also showed that items consisting of positive statements are skipped less often than items consisting of negative statements, $B = -1.20$, $z = -4.58$, $p < 0.001$. The level of openness of the respondent was positively associated with the tendency to refuse to choose one of the statements, $B = 0.30$, $z = 3.26$, $p < 0.01$. Consistent with the results of Model 3, the valence of the item did not moderate the association between openness and the propensity to refuse to choose an option, $B = -0.04$, $z = -1.13$, $p = 0.26$. Furthermore, consistent with Model 3, we found that the tendency to skip negative items more often than positive items appeared to be stronger among respondents who score high on agreeableness, $B = -0.13$, $z = -3.15$, $p < 0.01$, and emotional stability, $B = -0.27$, $z = -7.11$, $p < 0.001$, and marginally stronger among respondents who score high on extroversion, $B = -0.07$, $z = -1.85$, $p = 0.06$. The results of this simplified analysis essentially indicate that the introduction of Node 2 in Model 3 did not alter the associations between personality traits and the propensity to either provide or refuse to provide an answer (corresponding to Node 1 in Model 3).

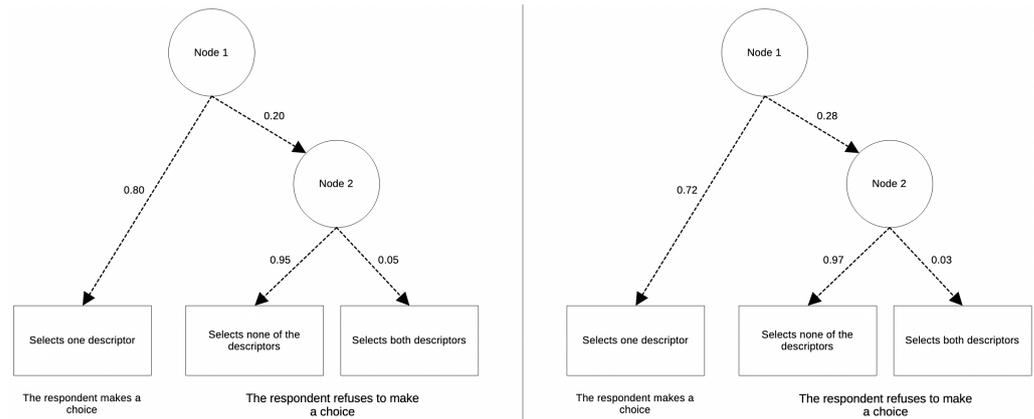


Figure 3. Item skipping and given reason at different levels of agreeableness for negative items. **(Left panel):** The respondent scores low on agreeableness (-1 SD). **(Right panel):** The respondent scores high on agreeableness ($+1$ SD).

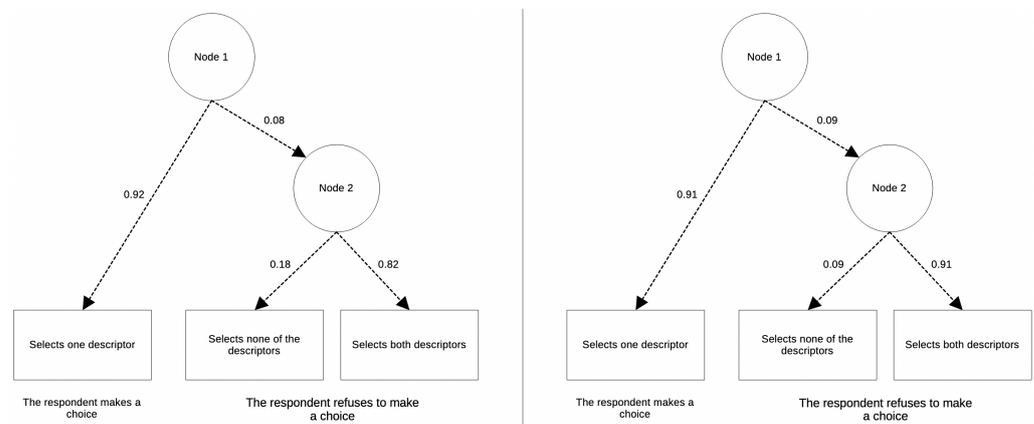


Figure 4. Item skipping and given reason at different levels of agreeableness for positive items. **(Left panel):** The respondent scores low on agreeableness (-1 SD). **(Right panel):** The respondent scores high on agreeableness ($+1$ SD).

4. Discussion

Our aim with the present study was to bridge the gap in the literature about the antecedents of item skipping when responding to forced-choice personality questionnaires. Based on two potential mechanisms to explain individual differences in item skipping in this context, we proposed hypotheses on the links between general personality traits and the propensity to skip items. On the one hand, we proposed that openness is positively associated with item skipping because it is associated with more complex self-schemas and the evaluation of several statements as equally descriptive of one's personality. On the other hand, we proposed that the traits associated with social desirability—that is, honesty/humility, emotional stability, extroversion, agreeableness, and conscientiousness [14]—are positively related to item skipping but mostly when the items consist of socially undesirable item statements.

Altogether, our findings suggest that different personality traits relate to item skipping through different mechanisms. For example, openness is associated with a general tendency to skip items, which we explain by open respondents identifying strongly with more personality descriptors than less open respondents. This is consistent with the idea that open respondents have more complex self-schemas, which leads them to skip items more often. Agreeableness is associated with the tendency skip negative items especially, possibly for the reason that respondents who are more agreeable do not identify with any negative personality descriptor in comparison with respondents who are less agreeable. This is

consistent with the idea that agreeable respondents have a higher need for social desirability, which leads them to more often skip negative items in particular.

4.1. Implications

Our findings suggest that the propensity to skip forced-choice personality items is not completely random, as it is associated with both person and item characteristics. Researchers or practitioners who administer forced-choice questionnaires and allow respondents to skip items should be aware that it is a potential source of bias that should be taken into consideration. One seemingly easy way to circumvent the psychometric problems caused by item skipping would be to simply remove the possibility to skip an item in the first place. This solution is relatively easy to implement, especially in online assessments in which forcing respondents to give a response before moving to the next item is technically feasible. However, allowing respondents to skip items aligns with the ethical principles of respecting individual autonomy in personality assessments [7] and also prevents respondents from either stopping prematurely or responding carelessly to the rest of the test. For these reasons, forcing respondents to make a choice may not be optimal either.

Instead, we should be trying to find ways of limiting the potentially deleterious effects of item skipping on personality assessments or even using the phenomenon to improve the assessment. One of the main implications of our findings is that certain personality traits are associated with a higher tendency to refuse to answer. Concretely, it means that respondents with high scores on these traits will have more missing responses compared to others. For these respondents, information is lost and the scores will tend to be less reliable. One solution to mitigate this problem is to use Computerized Adaptive Testing (CAT) in conjunction with a large item bank. With CAT, items are presented until (among other possible criteria) enough information is obtained about a person's location. With CAT, respondents who tend to skip more often would need to be exposed to more items before a reliable estimate of their personality traits is obtained, but the estimate would be ultimately as reliable as the estimate of respondents who tend to skip less items.

Furthermore, using the same IRT framework as the basis of the CAT, it is even theoretically possible to extract information about the respondents' personalities by analyzing their tendency to skip forced-choice items [28]. In other words, skipping would itself provide collateral information about the trait(s) measured. Based on our results, information about skipping items (and the reasons given for skipping the item) indirectly provides information about the complexity of the respondents' self-schemas and level of social desirability, and it directly provides information about their personality traits. The propensity to skip forced-choice items could therefore not only be seen as a threat, but also as an opportunity in psychological assessment.

4.2. Limitations and Future Research

To the best of our knowledge, our study is one of the first to investigate the antecedents of forced-choice item skipping. However, as any exploratory work, it has limitations. A first limitation is that we based our hypotheses on assumptions about psychological mechanisms that might lead respondents to avoid giving a response to items for one reason or another without directly testing these mechanisms. More studies are needed to test whether the complexity of self-schemas and social desirability are indeed the psychological mechanisms that explain the links we found between general personality traits and the propensity to skip items. A better understanding of the mechanisms would be useful not only to verify the relevance of our claims, but especially to potentially identify other antecedents of item skipping. Discovering more antecedents would mean that more information about the respondents could be inferred from their propensity to refuse to give responses to forced-choice items.

Additionally, we only explored interactions between item desirability (positive vs. negative) and the personality traits of respondents in predicting the refusal to choose.

Exploring interactions with other item characteristics could provide further insights into the phenomenon of item skipping. One promising idea is to investigate interactions between the personality dimensions assessed by the forced-choice item and the personality traits of the respondent when predicting item skipping. Previous research on latency and personality traits suggests that when respondents score extremely (low or high) on a trait, they provide faster answers because their self-schema is clearer [29,30]. This could mean that, for forced-choice items, respondents who score extremely high or low on a trait would be less likely to skip an item that relates to that trait. Examining these interactions could allow test administrators to extract additional information from item skipping and inform the design of more effective personality assessments.

Finally, the development of an integrated IRT model that combines Thurstonian IRT for forced-choice items and IRTree for modeling the skipping process, along with the reasons given by respondents, represents a valuable endeavor in the field of personality assessment. To the best of our knowledge, there is currently no single model that can comprehensively account for both the act of responding to a forced-choice item and the complex process of refusing to respond. Such a unified model would not only address the limitations of the model we used in our study—that is, a model that focuses on item skipping and not on the forced-choice—but would also open up new avenues for exploring the complex interplay between personality traits, item characteristics, and possibly the underlying mechanisms involved in giving a response to forced-choice items.

Author Contributions: Conceptualization, M.S., N.M., E.K., and S.B.; methodology, M.S. and N.M.; formal analysis, M.S.; data curation, E.K. and M.S.; writing—original draft preparation, M.S. and N.M.; writing—review and editing, M.S., N.M., E.K. and S.B.. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The dataset and the R code are available on https://osf.io/7bs8a/?view_only=a5a2cfe70950457ebe669b3db2fd5874 (accessed on 7 January 2024).

Conflicts of Interest: E.K. and S.B. hold positions in the company that owns the forced-choice personality questionnaire used in our study.

References

1. Brown, A.; Maydeu-Olivares, A. Item response modeling of forced-choice questionnaires. *Educ. Psychol. Meas.* **2011**, *71*, 460–502. [CrossRef]
2. Bürkner, P.C.; Schulte, N.; Holling, H. On the statistical and practical limitations of Thurstonian IRT models. *Educ. Psychol. Meas.* **2019**, *79*, 827–854. [CrossRef] [PubMed]
3. Watrin, L.; Geiger, M.; Spengler, M.; Wilhelm, O. Forced-choice versus Likert responses on an occupational Big Five questionnaire. *J. Individ. Differ.* **2019**, *40*, 134. [CrossRef]
4. Cao, M.; Drasgow, F. Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *J. Appl. Psychol.* **2019**, *104*, 1347. [CrossRef]
5. Speer, A.B.; Wegmeyer, L.J.; Tenbrink, A.P.; Delacruz, A.Y.; Christiansen, N.D.; Salim, R.M. Comparing forced-choice and single-stimulus personality scores on a level playing field: A meta-analysis of psychometric properties and susceptibility to faking. *J. Appl. Psychol.* **2023**. [CrossRef]
6. Zhang, B.; Sun, T.; Drasgow, F.; Chernyshenko, O.S.; Nye, C.D.; Stark, S.; White, L.A. Though forced, still valid: Psychometric equivalence of forced-choice and single-statement measures. *Organ. Res. Methods* **2020**, *23*, 569–590. [CrossRef]
7. Brabender, V.; Bricklin, P. Ethical issues in psychological assessment in different settings. *J. Personal. Assess.* **2001**, *77*, 192–194. [CrossRef]
8. Thurstone, L.L. A law of comparative judgment. *Psychol. Rev.* **1994**, *101*, 266. [CrossRef]
9. Linville, P.W. Self-complexity as a cognitive buffer against stress-related illness and depression. *J. Personal. Soc. Psychol.* **1987**, *52*, 663. [CrossRef]
10. Sadowski, C.J.; Cogburn, H.E. Need for cognition in the big-five factor structure. *J. Psychol.* **1997**, *131*, 307–312. [CrossRef]

11. Furnham, A.; Marks, J. Tolerance of ambiguity: A review of the recent literature. *Psychology* **2013**, *4*, 717–728. [[CrossRef](#)]
12. Tetlock, P.E.; Peterson, R.S.; Berry, J.M. Flattering and unflattering personality portraits of integratively simple and complex managers. *J. Personal. Soc. Psychol.* **1993**, *64*, 500. [[CrossRef](#)]
13. Lindeman, M.; Verkasalo, M. Personality, situation, and positive–negative asymmetry in socially desirable responding. *Eur. J. Personal.* **1995**, *9*, 125–134. [[CrossRef](#)]
14. de Vries, R.E.; Zettler, I.; Hilbig, B.E. Rethinking trait conceptions of social desirability scales: Impression management as an expression of honesty-humility. *Assessment* **2014**, *21*, 286–299. [[CrossRef](#)]
15. Brown, A.; Maydeu-Olivares, A. Fitting a Thurstonian IRT model to forced-choice data using Mplus. *Behav. Res. Methods* **2012**, *44*, 1135–1147. [[CrossRef](#)] [[PubMed](#)]
16. Jeon, M.; De Boeck, P. A generalized item response tree model for psychological assessments. *Behav. Res. Methods* **2016**, *48*, 1070–1085. [[CrossRef](#)]
17. Plieninger, H. Developing and applying IR-tree models: Guidelines, caveats, and an extension to multiple groups. *Organ. Res. Methods* **2021**, *24*, 654–670. [[CrossRef](#)]
18. Hommel, B.E. Expanding the methodological toolbox: Machine-based item desirability ratings as an alternative to human-based ratings. *Personal. Individ. Differ.* **2023**, *213*, 112307. [[CrossRef](#)]
19. Denissen, J.J.; Soto, C.J.; Geenen, R.; John, O.P.; van Aken, M.A. Incorporating prosocial vs. antisocial trait content in Big Five measurement: Lessons from the Big Five Inventory-2 (BFI-2). *J. Res. Personal.* **2022**, *96*, 104147. [[CrossRef](#)]
20. Lignier, B.; Petot, J.M.; Canada, B.; De Oliveira, P.; Nicolas, M.; Courtois, R.; John, O.; Plaisant, O.; Soto, C. Factor structure, psychometric properties, and validity of the Big Five Inventory-2 facets: Evidence from the French adaptation (BFI-2-Fr). *Curr. Psychol.* **2022**, *42*, 26099–26114. [[CrossRef](#)]
21. De Boeck, P.; Partchev, I. IRTrees: Tree-based item response models of the GLMM family. *J. Stat. Softw.* **2012**, *48*, 1–28. [[CrossRef](#)]
22. De Boeck, P. Random Item IRT Models. *Psychometrika* **2008**, *73*, 533–559. [[CrossRef](#)]
23. Bartoń, K. *MuMIn: Multi-Model Inference*, R Package Version 1.47.5; R Core Team: Vienna, Austria, 2023.
24. Nakagawa, S.; Johnson, P.C.; Schielzeth, H. The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface* **2017**, *14*, 20170213. [[CrossRef](#)] [[PubMed](#)]
25. Bates, D.; Mächler, M.; Bolker, B.M.; Walker, S.C. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [[CrossRef](#)]
26. Kuznetsova, A.; Brockhoff, P.B.; Christensen, R.H.B. lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **2017**, *82*, 1–26. [[CrossRef](#)]
27. De Ayala, R.J. *The Theory and Practice of Item Response Theory*, 2nd ed.; The Guilford Press: New York, NY, USA, 2022.
28. Maij-de Meij, A.M.; Kelderman, H.; van der Flier, H. Fitting a mixture item response theory model to personality questionnaire data: Characterizing latent classes and investigating possibilities for improving prediction. *Appl. Psychol. Meas.* **2008**, *32*, 611–631. [[CrossRef](#)]
29. Akrami, N.; Hedlund, L.E.; Ekehammar, B. Personality scale response latencies as self-schema indicators: The inverted-U effect revisited. *Personal. Individ. Differ.* **2007**, *43*, 611–618. [[CrossRef](#)]
30. Ranger, J.; Ortner, T.M. Assessing personality traits through response latencies using item response theory. *Educ. Psychol. Meas.* **2011**, *71*, 389–406. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.