Food Quality and Preference 135 (2026) 105691

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Comparing a multi-item Rasch scale for purchase intentions with preference elicitation using experimental auctions for consumer acceptance of biofortified food: An in-store study in Rwanda

Carl-Johan Lagerkvist [a,*], Julius J. Okello [b], Sylvester Ojwang [c], Herve Mwizerwa [d], Jean Claude Nshimiyimana [d], Jean Ndirigwe [e], John Bosco Shingiro [f], Jan Low [c]

[a] Swedish University of Agricultural Sciences, Department of Economics, PO Box 7013, 750 07 Uppsala, Sweden
[b] International Potato Center, Uganda
[c] International Potato Center, Nairobi, Kenya
[d] International Potato Center, Kigali, Rwanda
[e] Jean Ndirigwe: Rwanda Agricultural Board (RAB), Kigali, Rwanda
[f] John Bosco Shingiro, World Food Program, Kigali, Rwanda

ARTICLE INFO

Keywords:
Consumer research
Food biofortification
Multi-response methodology
Becker-DeGroot-Marschak (BDM)
Nutrition
Sensory product perceptions

ABSTRACT

Consumer acceptance of biofortified vitamin-A-rich products by urban populations in developing countries is an important preventive pathway for addressing vitamin A deficiencies by easing nutritional availability and access. The present study used data from a multi-variable in-store consumer study in Rwanda with bread and a snack product (mandazi) to develop and test a measure for stated purchase intentions based on Rasch scaling, and investigates the structural relationship between purchase intentions and multiple measures of potential drivers for purchase intentions. The study investigates the extent to which the purchase intent influence the decisions and monetary product preferences within an in-store Becker-DeGroot-Marschak (BDM) auction based on a product exchange to the biofortified product using the self-selected non-fortified product as reference. The key finding is that the nature of the relationship between the methods to elicit preferences is product-specific because of differences in the outcome of the BDM. The nature of the relationship also depends on whether the measurement uncertainty in the Rasch score for person locations for purchase intentions is considered. Including the measurement uncertainty reverses the extent to which the choice to bid and the bid amount within the BDM are driven by separate processes. Moreover, while actual liking helped predict purchase intentions, no evidence was found of such association for nutritional beliefs. Sensory attributes for the vitamin-A-biofortified products do not serve as key product features to promote (or detract from) purchase intentions. Overall, these findings indicate that auction-based methods and Rasch scales for purchase intentions are complementary measures in product research.

1. Introduction

Vitamin A deficiency (VAD) among infants, young children, and pregnant and breastfeeding mothers is a major public health problem in developing countries (World Health Organization (WHO), 2019). Rwanda has one of the highest incidences of VAD and had the lowest observed change from 1990 to 2019 (Zhao et al., 2022). In the past two decades, governments and non-governmental organizations in the south-Saharan region have promoted the adoption and consumption of vitamin-A-rich orange-fleshed sweetpotato (OFSP) varieties by farm households that serve as producers and consumers.

To stimulate early prevention and further adoption among urban non-farm households, there has been a on making vitamin-A-fortified food products more readily available by commercializing OFSP value chains. Targeting urban populations with biofortified vitamin-A-rich processed products to ease nutritional availability and access can complement other current efforts to address VAD; namely, supplementation and industrial fortification. Providing OFSP puree as a substitute for

* Corresponding author.
E-mail addresses: carl-johan.lagerkvist@slu.se (C.-J. Lagerkvist), j.ojwang@cgiar.org (S. Ojwang), jnshimiyimana@cgiar.org (J.C. Nshimiyimana), j.low@cgiar.org (J. Low).

https://doi.org/10.1016/j.foodqual.2025.105691
Received 19 December 2024; Received in revised form 11 August 2025; Accepted 4 September 2025
Available online 12 September 2025

wheat flour in bread has been found to match well with consumer sensory and hedonic preferences and has therefore been considered as a potential pathway to achieving increased vitamin A consumption (Lagerkvist et al., 2021; Low & Van Jaarsveld, 2008; Okello et al., 2021). Other studies have shown that domestic production, product sensory characteristics and provision of nutritional information is vital for consumer acceptance of vitamin-A-biofortified crops and products thereof (e.g., Muzhingi et al., 2008, Bocher et al., 2019).

Exploring the potential for marketing a wider array of products based on OFSP can direct interventions to positively impact vulnerable households and serve as a pathway to tackle VAD. To this end, hedonic, sensory and emotional evaluations have been widely used to predict food choices (e.g., Dalenberg et al., 2014; Forde & de Graaf, 2023; Mustonen et al., 2007) and sensory-driven linkages to product perceptions are well-supported (e.g., Giacalone & Jaeger, 2021). Furthermore, consumers' self-reported (that is, stated) preferences in terms of behavioral (purchase) intentions to predict choice, as well as hypothetical willingness-to-pay (HWTP), as a monetary value expression, are of core interest to food consumer researchers, marketers and the like. Although expressions of HWTP is comparable to behavioral intentions (Ajzen & Driver, 1992), the inconsistency (or, hypothetical) bias arising from people saying one thing but doing another is well documented (e.g. DeFleur & Westie, 1958). Still, these measures serve as a basis for marketing programs (e.g. Barber et al., 2012; Jamieson & Bass, 1989) and have been combined in commercial research with purchase intentions to forecast sales and make demand predictions (e.g. Armstrong et al., 2000).

Many previous studies on consumer stated purchase intentions have been based on psychology-based theories of attitude-behavior relations, such as within the theory of planned behavior (Ajzen, 1991). Here, an attitude is an independent concept that needs to become activated before it may, through intentions, trigger a certain behavior (e.g. Fazio et al., 2008). The empirical discrepancy between intentions and behavior is then attributed to either a disparity in beliefs activated in the behavioral versus the symbolic situation (Ajzen et al., 2004) or, as noted by Campbell (1963), from the disregard in the hypothetical situation from the relative situational difficulties of performing the behavior. The alternative theoretical approach is based on Campbell's (1963) conception that stated claims and behavior towards an attitude object are inseparable expressions of the same latent disposition and Greve's (2001) notion of a formal attitude-behavior relationship (that is, as an entity). Here, Kaiser, Byrka, and Hartig (2010) developed a paradigm for attitude research (henceforth, Campbell's paradigm) treating "individual behavior as a function of a person's attitude and of the difficulty of performing the given behavior" (p. 353). In their review of existing research, Kaiser et al. (2010) corroborated this means-end relationship describing an individual's attitude as the likelihood of engaging in behaviors with the specific domain. The existing literature exploring the approach by Kaiser et al. (2010) with applications to food consumer behavior is limited to work by Fischer and co-authors (Fischer et al., 2006; Fischer & Frewer, 2009) for handling of food safety risks. The approaches of Campbell (1963) and Kaiser et al. (2010) are relevant to behavioral and marketing intervention studies because the question of how to change the behavior becomes equivalent to the question of how to change the attitude.

In parallel, to overcome the issue of the hypothetical bias in economic preference elicitation, a growing body of literature has adopted auction-based methods to elicit real willingness to pay (RWTP). The monetary expression is then consequential, with real money being exchanged for an actual exchange of real products, as well as being set to be incentive-compatible (that is, participants have incentives to behave as they want to) (Caputo et al., 2023; Shogren et al., 2001). Auction-based valuation methods have become increasingly common in food and consumer research and are distinguishable in terms of how they are contextualized (from a lab to a natural environment) (see the review by Vecchio & Borrello, 2019). However, the reliability of laboratory and online auctions for elicitation of preferences has been questioned due to its lack of realism, bias in selection of participants, and for issues of strategic bias that occur if participants deliberately misrepresent their preferences (Bennett & Blamey, 2001). The most common methods used in field settings are the Vickery second price auction (Vickery, 1961) and the Becker-DeGroot-Marschak (BDM) mechanism (Becker et al., 1964). The BDM provides an incentive-compatible and demand-revealing measure of RWTP where each participant first decides whether or not to bid and then decides on the amount to bid. This bid is then compared to a randomly drawn price to determine whether the person will receive the item/product being auctioned and pay according to his/her bid (Lusk & Shogren, 2007).

Existing research examining the viability of economic and psychological perspectives on valuation has documented that HTWP represents evaluative expressions of attitudes rather than economic preferences (Kahneman et al., 1999; Kahneman & Ritov, 1994). To the best of our knowledge, within this line of attitude-behavior research contrasting psychology and economics perspectives no previous study has compared the conceptualization of attitudes as behavioral dispositions based on Campbell's paradigm and the economically consequential behavior expressed through auction-based methods. Therefore, the main objective of the present study is to evaluate the relationship between consumers' purchase intentions and their RWTP when matching vitamin A enrichment based on OFSP to new products in demand among urban populations in Kigali, Uganda. To this end, the present study: (i) develops a measure for purchase intentions based on Campbell's paradigm, which serves to distinguish the contribution of product item characteristics and person behavioral difficulties to the formation of purchase intentions; (ii) investigates the structural relationship between the purchase intention measure and multiple measures of potential drivers for the purchase intention; and (iii) investigates the extent to which purchase intent influences decisions made within the in-store auction in terms of product exchange, and also whether the monetary preference in such a case is influenced by the decision to be willing to exchange.

2. Theoretical background and the present study

There are five ways in which this study differs from previous research on consumers' stated purchase intentions and valuation of food products. First, existing research examining consumers stated purchase intentions and actual purchase behavior has been restricted to the investigation of one type of product (e.g., Barber et al., 2012). We included two OFSP-based products, which address the extent of product uniqueness for the relationship between stated purchase intentions and revealed preference.

Second, within studies based on psychology-based theories of attitude-behavior relations, the intention construct typically refers to "instructions that people give to themselves to behave in certain ways" (Triandis, 1980, p.203) and indexes a mental readiness to perform the behavior ("I intend to do X") based on an evaluative tendency to an object in combination with subjective norm and perception of behavioral control (Eagly & Chaiken, 1993; Sheeran, 2002). Instead, we adopted the implementation of Campbell's paradigm by Kaiser et al. (2010) and develop product-specific scales for stated purchase intentions for the two OFSP-based products. This allowed us to evaluate how a set of key product feature characteristics contributes to the purchase intention and whether there are product-specific differences in key factors for successfully introducing vitamin-A-fortified to the targeted population. With this approach, the purchase disposition is conceptualized through behavioral means by which consumers manifest their specific attitude. Here, a set of behavioral performances, such as helping one's children to eat healthy, supporting local production of ingredients, and seeking tastiness, constitute a uniform dimension of behaviors. Along this dimension, the behavioral means are ordered transitively in terms of their difficulty representing different levels of the attitudinal goal. An assumption of uniform effective attitudes is made so

that the attitude effectively serves as motivation along the full range of behaviors regardless of the difficulty of the behavioral means. Furthermore, and importantly, behavioral difficulties represent situational forces that are decisive for performing certain behaviors. These difficulties are assumed to be external and therefore "operate independent of the actors involved, regardless of personal attitude levels, regardless of the perception of obstacles or facilitators, and regardless of any personal differences in the capability to perform the behavior" (Kaiser et al., 2010, p. 359). Accordingly, the difficulty of a behavioral means is recognized through the proportion of individuals who enact it.

Third, results from existing research based on the classic mental readiness approach support the argument that sociodemographic factors, as well as cognitive and behavioral factors, contribute to explain purchase intentions for vitamin-A-biofortified rice (e.g. Ethen et al., 2024). In a broader context of food choice, the extent to which nutritional beliefs influence choice draws on early findings by Logue (1986) but a recent extensive review by Ballco and Gracia (2022) reported mixed evidence on consumer preferences. However, Naico and Lusk (2010) found that the willingness to consume OFSP instead of traditional varieties is influenced by information about nutritional benefits. In Lagerkvist et al.'s (2021)in-store auction study on consumers' real willingness to pay for bread based on OFSP puree in Kenya, nutritional beliefs influenced the evaluation of sensory attributes, as well as the liking assessment, but were not predictive of actual choice of the OFSP product. However, the liking assessment was strongly predictive of choice. For food products, the predictive accuracy of purchase intentions to purchase behavior has been shown to increase when participants have been able to taste the product before rating their purchase intentions (Kytö et al., 2019). The present study contributes to this topic and examines whether sociodemographic factors and nutritional beliefs as well as liking after tasting predict consumers' purchase intentions or choice, or both.

Fourth, few studies have sought to understand consumer preferences for enhanced levels of vitamin-A in a developing country context using the BDM mechanism. Results based on home-use tests show a potential market for fortified cereal food products and indicate that consumer sensory acceptance and information about nutritional benefits serve to relate to the RWTP (e.g. De Groote et al., 2014). While there have been mixed results regarding how personal characteristics relate to RWTP, research findings do support higher acceptance in higher income and education groups (e.g. De Groote et al., 2018). More recently, auction-based research to evaluate food consumers has revealed that monetary product preferences have moved towards the in-store environment to mimic the consumer product choice decision process as naturally as possible (Xue, Mainville, You and Nayga Jr., 2010; Lagerkvist & Okello, 2016; Carrol & Samek, 2018; Lagerkvist et al., 2021). Lagerkvist et al.'s (2021) in-store auction study on consumers' RWTP for bread based on OFSP puree in Kenya by, compared WTP between two buyer categories (that is, those initially selecting OFSP bread and those selecting other bread). By contrast, the present study was designed to mimic the consumer product choice decision process in a market where the new product is not yet available. As with Treatment B in Lagerkvist et al. (2021), participants were recruited after having been observed to select an existing product within the product category for actual purchase. Thus, participants are self-revealed as prospective buyers. Although the current study design did not include a baseline assessment of the initially selected product, the self-selected product served as reference point of each participant when reporting on their value of the OFSP product. This approach directs attention on the difference in WTP between the two products and has been reported to provide the most reliable estimates of WTP from experimental auctions (Hoffman et al. (1993). Furthermore, and as an extension of Lagerkvist et al. (2021), for participants who were not willing to provide a bid to replace the initially selected product with the OFSP alternative, a follow-up question was used to identify the reason for this decision. If the reason was related to price (that is, not being willing to pay for an exchange), we instead elicited the willingness-to-accept (WTA) (e.g., Irwin et al., 1998; Marzilli Ericson & Fuster, 2011). The WTA represents the price discount at which the respondent would be willing to trade off the status quo product for the OFSP alternative. This two-sided (WTP and WTA) approach meant that each participant was free to act according to his or her own preferences.

Fifth, we examined the relationship between the product-specific scales for stated purchase intentions and each of the two modes of valuation within the BDM mechanism. Kahneman et al. (1999) support the argument that differences in valuation are found between monetary measures and in ratings (or based on attitudes). The first step in the BDM represents a judgement about participation choice about committing to bid (yes/no) for a possible exchange of two products. This is similar to expressing a stated purchase intention but goes beyond a mere intention as it is consequential. A 'yes' is a commitment to pay for the exchange or forego the chance of exchange depending on the outcome of the random draw for the reservation price in the second step of the BDM. In case of a 'yes' to bid, the monetary valuation in the second part in the BDM entails a decision about the amount of the actual bid ("how much"); that is, the reservation price for the exchange. A novel contribution of the present study is that we examined whether the purchase intention measure influences the two BDM decisions differently. Moreover, we examined the extent to which the choice to bid and the actual money bid were interlinked. The extent to which bids are influenced by the decision to participate represents a potential bias of the BDM mechanism, which, to our knowledge, has not been addressed in previous research. For product-based research, evidence of bids being inflated by the decision to bid would challenge the reliability of reservation prices elicited by the BDM mechanism. Furthermore, the measure of purchase intention (see Section 3.5) provides estimates of measurement uncertainty for each individual's scale location. Therefore, a specific contribution of the present study is to assess the extent to which this measurement uncertainty contributes to the potential discrepancy between stated purchase intentions and RWTP.

3. Materials and methods

This study was conducted according to the guidelines set out in the Declaration of Helsinki and all procedures involving human subjects were approved by the National Institute of Statistics in Rwanda on Dec 2, 2022 (No. 0532/2022/10/INSR).

3.1. Case study: Bread and mandazi baked using orange-fleshed sweetpotato puree

Bread and donuts are widely consumed as breakfast and snack meals in urban areas in Rwanda, which potentially makes them good vehicles for providing vitamin A to urban consumers. The present study focused on initiatives from private businesses to market donuts and bread baked using OFSP puree. OFSP mandazi is a fried product (similar to a donut) that is displayed non-packed in open baskets in shops. OFSP-based donuts were first introduced in Rwanda targeting rural consumers, but regular production failed to pick up due to organizational/management issues (Muoki & Kwikiriza, 2019). The OFSP-based bread (named VITA bread) is the latest addition to the market and is primarily sold in metropolitan Kigali. The VITA bread is baked and displayed in branded packages, just like other bread brands on supermarket shelves.

3.2. Participants

After a set of focus group discussions and pre-tests that had the aim of adapting the study instrument to the local context, an in-store field experiment on bread and mandazi choices was conducted in December 2022 in eight different supermarkets/stores (for the VITA bread), and seven Sina Gerard stores (the only outlets for mandazi) in the Kigali City province (Rwanda). The stores were selected from all three districts of the province.

Enumerators invited every second shopper they observed selecting the non-OFSP bread/mandazi to participate in the study, until a sample size of at least 150 people for bread and 200 for mandazi had been achieved. Following Lagerkvist et al. (2021), recruitment was done after a shopper had been observed taking the product from the shelf and was about to leave the bread/mandazi section of the store. Each prospective respondent was informed about the objectives of the research study and then asked for their consent to participate in the study. Interviews were conducted in-store if the response was affirmative. To capture different shopping patterns, interviews were conducted between 7:00 a.m. and 8:00 p.m. on every day of the week except Sundays. Sunday was excluded because most city stores do not receive a significant number of shoppers given that most people stay home and/or go to church. Cross-participation was controlled for by telephone numbers and a screening question that enabled the enumerators to determine whether the respondent has been interviewed elsewhere. A total of 168 interviews were conducted for bread and 218 for mandazi.

3.3. Steps in the field study

Eight trained enumerators conducted the study using tablet-based questionnaires. The study was conducted either in the local Kinyarwanda language or in English, depending on the consumer's preference.

In Step 1, participants were provided with a brief product narrative for the VITA bread/OFSP mandazi (Appendix 1) and then invited to taste a product sample. Prior to tasting, participants were asked to rinse their mouth with still bottled water. In Step 2, participants were instructed to rate their nutritional beliefs, as well as how much they liked the product. In Step 3, participants were invited to taste the product again and then asked to rate statements referring to purchase intentions (20 for VITA bread, 19 for OFSP mandazi, Appendix 2). The consumption requirement was expected to increase the attentiveness and the preparedness leading into the evaluation tasks.

In Step 4 (see Appendix 3 for the detailed instructions), participants were asked to present the bread/mandazi they had initially selected to buy and were reminded that they had just tasted the VITA bread/OFSP mandazi. Participants then received a participation fee of 2000RWF (≈USD 1.83) for the bread study and 500RWF (USD 0.46) for the mandazi study. Next, participants were asked whether or not they would be willing to use some portion of the money they had just received to exchange the initially selected product for the VITA bread/OFSP mandazi (see Supplementary Material for instructions). Each participant was informed that, depending on his or her choice of the extra amount that he or she was willing to provide for the exchange and with a randomly drawn strike price, they might instead end up paying for the VITA bread/OFSP mandazi and leave the initially selected product at the store. If the answer was affirmative, the BDM mechanism (Becker, DeGroot, Marschak, 1964) was used to elicit the participant's WTP for the exchange. Before placing their bids, participants were provided with an example of the BDM mechanism.

If a participant declined to bid for the exchange, he or she was asked a follow-up question about the reasons for this decision. If the reason was price, participants were asked about their interest in leaving with the VITA bread/OFSP mandazi for a certain compensation. This was intended to mimic a discount that would be sought in order to exchange the initially selected product. If this interest was confirmed, a second BDM was used to elicit the participant's WTA the exchange. Before placing their bids for compensation, participants were provided with an example of the BDM mechanism adopted for the purpose of eliciting WTA. However, if a participant was unwilling to receive a potential compensation, he or she was asked to choose from a set of nine alternatives on a scale from 1 (strongly disagree) to 5 (strongly agree) regarding the reasons for this decision.

After the BDM, a post-auction survey module was administered. The survey collected data about socio-demographic characteristics and weekly consumption frequency for bread/mandazi.

3.4. Measures

3.4.1. Purchase intentions

Regarding the context of the current research, the acceptability of the VITA bread/OFSP mandazi was directed towards future behavior in terms of purchase intentions. The third stage of the survey consisted of 20 items regarding participants' intentions to purchase VITA bread/OFSP mandazi (see Appendix 2). The items were selected following consultation with scientists (agronomists, food scientist/technologist, and economists) who were familiar with OFSP production and processing. The items were tested with consumers in a separate focus group discussion. For mandazi, Item 12 was not included because mandazi is sold unpackaged. Responses were made on a five-point scale, ranging from 1 (completely disagree) to 5 (fully agree). Items were presented to participants in a randomized order.

3.4.2. Nutritional beliefs and liking after tasting

Following Lagerkvist et al. (2021), nutritional beliefs were assessed using a seven-point scale, ranging from −3 (not at all nutritious) to 3 (very nutritious). Actual liking was assessed using a seven-point hedonic scale ranging from −3 (dislike it very much) to 3 (like it very much).

3.5. Analyses

3.5.1. Rasch model for purchase intentions

We adopted Kaiser et al.'s (2010) conceptualization of Campbell's paradigm for modeling the behavioral disposition (an individual's purchase intention) as a function of the attitude and the difficulties (costs) for the set of behavioral items available (Appendix 2) for realizing the attitudinal goal. Notably, when consumers act to realize their purchase intentions through the defined set of behaviors, it is reasonable to expect that irregularities may matter due to individual differences and capabilities to perform the behaviors, as well as due to the situational implementation and selection among the behavioral means. Therefore, and following Kaiser et al. (2010), we examined the probability of the engagement that individuals expend to implement their attitude. Specifically, serving as a basis for the analysis, according to the Rasch (1980) model depicted in Eq. (1), the probability (that is, the odds) of engaging in a specific behavior associated with the purchase intention is detailed as the difference between the strength of the individual attitude and the difficulty of the behavior:

$$ln\left(\frac{p_{ki}}{1-p_{ki}}\right) = \theta_k - \delta_i \tag{1}$$

where p_{ki} depicts the probability of individual k's engagement in a specific behavior i with θ_k representing the individual attitude and δ_i represents the item-specific behavioral difficulty. According to Eq. (1), the attitude level is person-specific, whereas behaviors are characterized by how difficult they are to realize from a population perspective. The model in Eq. (1) yields estimates of person (attitude) and item (difficulty) locations that are independent from each other (Ekstrand, 2022).

Because purchase intentions were measured using ordered response categories, the scoring function in the polytomous Rasch model takes the following general form (Ekstrand, 2022):

$$P_{nix} = \frac{exp^{-\tau_{1i}-\tau_{2i}\ldots-\tau_{xi}.+x(\beta_n-\delta_i)}}{\sum_{x'}^{m_i} exp^{-\tau_{1i}-\tau_{2i}\ldots-\tau_{x'i}.+x'(\beta_n-\delta_i)}} \tag{2}$$

where P_{nix} is the probability that person n will score x on item i; for each item, τ_{xi} ($x = 1, 2,.., m_i$) are the thresholds that separate the latent continuum into $m + 1$ ordered categories.

The estimation of (2) provides linearization into interval logit measures on the negative-to-positive range. A location that is higher on the logit-transformed item range expresses a higher individual purchase intention likelihood based on the set of behaviors (items) that expresses the attitude. Furthermore, the Rasch model provides direct estimates of

the measurement uncertainty in its precision (standard error, SE) for both item and person locations on the latent continuum. We followed the approach developed by Ekstrand et al. (2022; Eq. 4 and Eq. 5) and transformed the logit locations into a 0–100 (interval) range, as well as in transforming the associated SEs to provide information on the measurement precision on the new transformed scale. This transform facilitates the further use in parametric regression analysis and eases comparison across products.

The analysis was performed using RUMM2030Plus (Version 5.8.1; Andrich et al., 2010) and we followed Tennant and Conaghan's (2007) recommended list of quality identifiers for reporting Rasch analysis. Specifically, the following steps were completed to obtain the person locations from the Rasch scale for each product, respectively. First, a likelihood ratio test was used to determine whether the Andrich (1978) rating scale (fixed thresholds for all items) or the Masters (1982) partial credit model was most appropriate for the polytomous data. Second, we checked the ordering of response categories and rescored by collapsing categories in order to eliminate disordered thresholds as much as possible. Third, we evaluated item fit using transformed residuals outside the ±2.5 range with Bonferroni base 0.05 adjusted probabilities to indicate misfit. Furthermore, item–trait interaction was evaluated because it provides a measure of all the item or person deviations from the Rasch model. In terms of item fit, the item–trait interaction reflects the property of invariance of the person ability to the hierarchical ordering of items by their difficulty. Fourth, local item independence was assessed by analyzing the correlation of unexplained variances for items with residuals less than 0.2 above the average residual correlation to be acceptable (Pallant & Tennant, 2007). We assessed unidimensionality in terms of measuring the latent purchase intention trait by paired *t*-tests (at $p < 0.05$) using the positively and negatively loaded items (threshold 0.3) on the first principal component. The Rasch scale is deemed to be unidimensional if the percentage of significant t-test is less than 5 % (Hagell, 2014). The fifth step was identifying the presence of differential item functioning (DIF) (item bias by subgroup differences in response to specific items). DIF was evaluated with a between-groups analysis of variance (ANOVA). The sixth step was targeting the scale in terms of mean location of persons and items (within ±0.5 indicate well-targeting). Finally, the person separation index (should be above 0.8 for individual assessment) and Cronbach's alpha were used as measures of reliability of the scale.

3.5.2. Conditional inference analysis for purchase intentions

The measurement invariance of the product specific purchase intention measures developed through the Rasch analysis was examined in terms of stability and sensitivity to persons and subgroups characteristics within the data available. In Rasch scale development, analysis of differential item functioning (DIF) where, say, men and women systematically score an item differently is relevant when examining potential presence of item bias (Tennant and Küçükdeveci, 2023). Because DIF examines the difference in the conditional scoring probabilities between groups of a certain variable, and with the grouping to be specified by the analyst, a key aspect of Rasch analysis is to set the grouping in accordance with those who possess the same underlying true ability. However, in the absence of pre-defined groups based on theory, the adequate grouping for continuous variables (like age) or for the multi-categorical variables (liking, income, etc.) to be used in DIF analysis is a daunting problem that requires extensive pre-testing. Therefore, including several potentially ad hoc groupings as basis for partial DIF analyses may compromise test ethics. Moreover, Tennant and Küçükdeveci (2023) noted that "contextual variables should not be on the (hypothesized) causal pathway… as their effect upon (the dependent variable) could be mediated, such that the effect of mediation could manifest as DIF" (p. 6).

Predictive modeling methods based on machine learning algorithms have become popular across various disciplines for subgroup analysis (e.g. Seibold et al., 2016). The subgroup identification seeks to explain differential outcomes for a dependent variable based on automated interaction detection from a set of predictive covariates. Therefore, instead of performing partial tests for DIF, we adopted the unbiased recursive partitioning model using conditional inference trees (ctree) (Hothorn et al., 2006). Based on ordinal regression, the ctree model was set to analyze the variable selection and the explanatory power of the purchase intention to the ratings of nutritional beliefs, actual liking and a set of socio-demographic variables as covariates, including weekly consumption frequency. The contextual variables that we included in the ctree analysis represent potential mediators (product type, nutritional beliefs, and actual liking), as well as potential covariates to such mediations. The approach of using the ctree algorithm (with Bonferroni adjustments) is not compromised in terms of multiple-hypotheses testing.

The ctree approach is based on conditional probability distributions between the purchase intention score (dependent variable) and the set of explanatory variables, and therefore matches with the test logic within DIF analysis for Rasch models. The 'ctree' utilizes statistical permutation tests to determine the optimal tree size, thereby overcoming a major limitation of from its 'predecessor' – classification and regression trees (CART; Breiman et al., 1984) – namely, the selection bias towards predictors with many potential split points. Furthermore, the ctree model can be used without transformation of continuous and ordinal exploratory variables. This is an advantage over the alternative CHAID model (Kass, 1980), which can only be applied when data is categorical in nature. Moreover, using the ctree approach overcomes the potential model specification bias when specifying a parametric regression model without selecting those variables and transformations thereof with the highest relevance, with the largest statistical explanatory power and including the full set of relationships among the covariates.

The ctree approach works in three steps to identify the structural relation among unbiased splits within the covariate space, for which the dependent variable is the most different. Initially, a global null hypothesis of no relationship between the covariates and the dependent variable is tested. If that hypothesis is rejected, the first binary split is for the covariate with the largest association with the dependent variable. The analysis is then re-iterated until global independence cannot be rejected at the pre-set significance level (base $\alpha = 0.05$, Bonferroni-adjusted).

To control for so-called pathological splits (Hothorn et al., 2006), a minprob criterion of 20 % was used. This establishes the proportion of observations needed to establish a terminal node. Furthermore, a minsplit criterion was set at 0.2 as the minimum sum of weights in a node to be considered for splitting to control for the proportion of observations needed to establish a terminal node. The ctree analyses were estimated in the partykit package (version 1.2–16) (Hothorn & Zeileis, 2015) for R 4.2.2 (R Core Team, 2022).

3.5.3. Examining the relationship between purchase intentions and decisions within the BDM

Some important analytical aspects in relation to the data generating process need to be considered when examining whether the purchase intention measure influences the two BDM decisions differently, and when examining whether the choice to bid and the actual money bid are interlinked.

When abstaining from the choice to bid, participants stay with the initially selected product, which means that they have indicated that they attribute a genuine zero utility preference for the exchange product (that is, bid = 0 also implies WTP = 0). The presence of so-called true (that is, by optimal choice) zero-bids introduces a bias when using ordinary least square (OLS) regression to estimate unknown parameters of a regression model. Therefore, to examine how the measure of purchase intention helps explain the decisions made in the BDM mechanism, we compared censored (tobit) with a double-hurdle model regression (Cragg, 1971). The key differences between these models concern the nature of the data censoring and whether the hole mixture (including the

zero-bidding participants) is relevant to consider in the analysis or, alternatively, whether the decisions in the field auction should be sought only for the subset of the samples that decided to participate.

If the purchase intent measure explains the decision about whether or not to bid (yes/no), and simultaneously, explains the level of the bid when provided (how much"), a censored tobit model (Tobin, 1958) would be appropriate (Wooldridge, 2002). In this case, the zero and non-zero WTPs come from the same data generating process and follow a mixed distribution with a probability mass at zero and a continuous distribution of WTP values above zero.

Alternatively, the choice to bid and the decision on the amount to bid can be made in two steps, or it is possible that the measure of purchase intent explains the decision to bid and the level of the bid differently. This can also entail the possibility of either strategic bids, or inflated bids, both of which represent situations where the level of the bid is influenced by the decision to bid. Cragg's model is then appropriate (Wooldridge, 2002). The first hurdle part fits a binomial (logit) regression model for the exceedance probability of the choice to bid. In the second part, a truncated regression model fits the WTP to its covariates for the subset of the data given by the first hurdle.

The tobit model was estimated using the 'survreg' function in the 'survival' package (v3.2–7; Therneau, 2020), while the first part of Cragg's model was estimated using the glm function with a logit link function and the second part was estimated using the 'truncreg' package (version 0.2–5; Croissant & Zeileis, 2018) for R 4.2.2 (R Core Team, 2022). For both models, we used the transformed Rasch score (Ekstrand et al., 2022) as the explanatory variable. For the comparison between the tobit model and Cragg's model, we used likelihood ratio tests (LRT).

Next, applications of censored and double-hurdle regression models to experimental auction data typically assume that the variance of the auction bids are constant, which means that the variance does not depend on the value of (any) covariates (that is, homoscedasticity). Empirically, this assumption can be challenged for reasons such as low sample sizes, or when the distribution of WTP is skewed, or in presence of outlier bids of the WTP, for example, due to more extreme bids to the end of the WTP distribution due to either protest(deflated) bids or when bids are overstated (inflated). A further source of non-constant variance to consider is when the distribution of a covariate is obtained through model estimation. Importantly, when estimating the Rasch model for purchase intentions, standard errors are provided as a measure of measurement uncertainty at each estimated mean person ability. This information may help to further explain how purchase intentions are related to the WTP. We adopted the conditional heteroscedastic censored regression approach of Messner et al. (2016) with the transformed person ability location (Rasch score) as regressor, and the log of the associated standard error as regressor for the scale, as follows:

$$wtp = \alpha + \beta(Rasch\ score_i)\,|\log(standard\ error\ Rasch\ score_i) + \varepsilon, wtp > 0 \quad (3)$$

where i denotes the i-th participant, and where α and ε denote the intercept and the composite error term, respectively. The β-coefficient represents the average relationship between the Rasch score and the WTP bid. We compared the fit of (1) to the nested constant zero-scale tobit model based on AIC and BIC.

We compared the censored conditional model in (1) with a two-part model double-hurdle conditional model consisting of: (a) a heteroscedastic logit model and (b) a truncated conditional regression model with the same specification as in Eq. (1). However, the estimation of the logit model does not identify the intercept of the scale model. Following Messner et al. (2016), the location coefficients of the tobit model and the truncated model were scaled to allow a comparison of fit using the AIC and BIC measures. The heteroscedastic censored and truncated regressions were performed using the package 'crch' package (version 1.2–1; Zeileis, Messner, et al., 2024), while the first part of the conditional hurdle model was estimated using the 'glmx' package (version 0.2–1; Zeileis, Koenker, & Doebler, 2024) for R 4.2.2 (R Core Team, 2022).

4. Results

4.1. Sample characteristics

White bread (127) dominated the initial selection over brown (29) and yellow (12) bread. The average weight was 657 g (stdev = 103) and the average price/kg was 2378RWF (stdev = 506). For mandazi, only one product was available, which was sold for 200RWF. Table 1 presents the descriptive socio-demographic statistics of the two samples. Summary statistics for nutritional beliefs, liking, and consumption frequency are presented in Table 2.

4.2. Rasch analysis of purchase intentions

Table 3 presents the overall Rasch fit statistics for the product-specific purchase intention scales. In the initial analysis for VITA bread, only four participants showed fit residuals outside the ±2.5 range. These individuals were excluded from the subsequent analysis because person misfits may cause deviation from model expectation at the item level (Pallant & Tennant, 2007). The initial analysis based on all items indicated the suitability of the partial credit model (χ^2 146.8 (df = 56), $p < 0.001$). The item–trait interaction (Bread A, 20 items) indicated a misfit to the model's expectations, as well as a slight mismatch between the person and item threshold distribution, with a positive mean person location indicating that a sample has higher purchase intentions than those captured by the average difficulty of the items. Local independence was supported for all item pairs except for Items 11 and 20 (corr = 0.367) and for Items 13 and 17 (corr = 0.572). The multidimensionality was not violated and no evidence of differences in relation to item difficulties by gender (DIF) was detected. The PSI measure of internal consistency at 0.86 was satisfactory (corresponding to Cronbach's alpha = 0.90), suggesting that the measurement scale was able to differentiate persons along the latent trait (Hagquist et al., 2009). Notably, Items 6, 15, and 19 displayed excessive fit residuals and Bonferroni adjusted p-values below 0.05/20. Moreover, significantly disordered thresholds were displayed for 18 items (that is, all items except Items 8 and 14). Although the findings of disordered thresholds do not violate the Rasch model, this issue may arise when response categories are observed relatively rarely (Tennant & Conaghan, 2007). The response pattern (see Appendix 2) showed few observations within the two categories indicating intentions not to be willing to buy (that is, Response Categories 1 and 2).

To resolve the issue of item and person fit, as well as to address the extent of disordered thresholds, a second model (Bread B) was re-analyzed without the three mis-fitting items and with use of four response categories for all items (Categories 1 and 2 collapsed). Six participants were excluded from the subsequent analysis based on fit residuals outside the ±2.5 range. A log-likelihood ratio test re-

Table 1
Demographics and summary statistics of the subsamples.

	Bread (n = 168)	Mandazi (n = 224)
Age (years)	31.1 (10.2)	30.2 (9.06)
Females:Males %	48.8:51.2	43.3:56.7
Years of education	11.7 (4.1)	10.6 (5.3)
Residence area (proportions):		
Urban	0.93	0.88
Peri-urban	0.07	0.12
Income/month (RWF):		
<50,000	0.28	0.35
50,000–200,000	0.48	0.53
201,000–500,000	0.14	0.09
>501,000	0.10	0.03
Consumption freq./week (average)	3.92	3.45

Note: Standard deviation in parenthesis.

Table 2
Summary statistics for liking and nutritional beliefs (proportions).

Liking	Dislike very much	Dislike moderately	Dislike slightly	Neither like nor dislike	Like slightly	Like Moderately	Like it very much
Bread	0.006	0.0119	0.0119	0.0595	0.1607	0.3929	0.3571
Mandazi	0	0.009	0.0089	0.0491	0.1473	0.3616	0.4241
Nutritional beliefs	Completely non-nutritious	Moderately non-nutritious	Non-nutritious	Neither non-nutritious nor nutritious	Nutritious	Moderately nutritious	Very nutritious
Bread	0	0.006	0.006	0.1071	0.1964	0.4702	0.2143
Mandazi	0	0.009	0	0.058	0.2589	0.4152	0.2589

Note: Based on complete samples (bread = 168, mandazi = 224).

Table 3
Rasch model summary statistics for overall fit of the product-specific scales for purchase intentions.

	Item		Person		Item-Trait interaction	PSI	Dimensionality	MIRC	DIF (gender)
Analysis	Location mean	Fit resid.	Location mean	Fit resid.	χ^2(df)/p-value		% t-tests >5 %		
Bread A	0.00	0.35	1.24	−0.29	128.3 (40)/<0.001	0.86	4.76 %	0.109	No
Bread B	0.00	0.11	1.08	−0.35	46.9(34)/0.07	0.85	4.17 %	0.052	No
Mandazi A	0.00	0.25	1.80	−0.23	97.7 (3)/<0.001	0.86	4.02 %	0.048	No
Mandazi B	0.00	0.04	1.42	−0.23	44.5 (32)/0.07	0.86	3.92 %	0.044	No

Note: df = degrees of freedom, PSI = person separation index, MIRC = mean item residual correlation, DIF = differential item functioning.

confirmed the appropriateness of the partial credit model after these adjustments (χ^2 78.8 (df = 31), $p < 0.001$). As Table 3 shows, the adjusted scale displayed acceptable fit in terms of item–trait interaction, PSI, and level of unidimensionality, as well as improvement in terms of targeting of locations for persons. The Cronbach's alpha for Bread B was 0.87. Furthermore, local independence was supported although item correlations between Items 11 and 20 and between Items 13 and 17. Person-item threshold distributions per product are shown in Appendix 4.

In the initial analysis for OFSP mandazi, six participants showed extreme fit residuals. The partial credit model was supported (χ^2 240.3 (df = 53), p < 0.001), but the item–trait interaction, as well as high person location mean, indicated a mismatch to the Rasch model. However, local independence was supported for all item pairs except for Items 12 and 16 (corr = 0.546). Furthermore, the model displayed good reliability (corresponding to Cronbach's alpha = 0.904), no evidence of subgroup differences by gender, and multidimensionality was not violated. Six items displayed disordered thresholds. Item 5 displayed extreme fit residuals (3.04, Bonferroni adjusted *p*-value = 0.0015). To resolve the issue of item fit and to address the extent of disordered thresholds, a second model was re-analyzed with Item 5 excluded and with use of four response categories for all items (Categories 1 and 2 collapsed). However, for the resulting model, the item–trait interaction still indicated an inadequate fit and multidimensionality was violated (8.93 % of *t*-tests <5 %). In the aggregate strategy to address the internal construct validity, we followed Pallant and Tennant (2007) and first examined whether we could identify individuals who had responded in a different way to the rest of the sample. Using Smith's Jr. (2002) *t*-test approach on the scoring on the first principal component, 20 individuals were identified with location differences outside the 95 % confidence interval (with 11 large negative and nine large positive sum scores). These individuals were predominantly male (65 %) with low income (65 %) with an average age (25.3) lower than the general average (30.7). For the resulting model with these respondents excluded, multidimensionality was not violated but the item–trait interaction still indicated a misfit to model expectations (χ^2 61.7 (df = 36), $p = 0.004$), with Items 11 and 13 contributing to this (Bonferroni adjusted *p*-values 0.009 and 0.018, respectively). The final scale with 16 items (Mandazi B) displayed good item fit and an improved person fit. A log-likelihood ratio test re-confirmed the appropriateness of the partial credit model after these adjustments (χ^2 85.5 (df = 29), p < 0.001).

The results for item locations for OFSP bread display an item ordering on a relatively narrow range of difficulties from the mid-point of the scale (logit = 0). The results suggest that creating jobs for youth and women (Item 4), high content of vitamin A (Item 20), nostalgia (Item 16) and preventing eyesight problems (Item 11) were the easiest items to endorse as a reason for purchase intentions (Table 4). On the other hand, packaging for freshness (Item 12), softness (Item 13), and crop draught tolerance (Item 8) were the most difficult items to endorse.

Furthermore, in relation to how each item contributes to the functioning of the purchase intention scale as a whole, there was a balance between the number of items undermining the total score (eight items with negative values for fit residuals) and over-discriminating items (eight items with positive values). Although the magnitudes of these residuals are not excessive, there are relatively more over-discriminating items at higher levels of item difficulty.

Items 6, 15, and 19 (see Appendix 2 for details) removed because of misfit. SE = standard error. df = 149.53 for fit residual.

The results for item locations for OFSP mandazi suggest a wider range of item difficulties for mandazi compared to the range for the VITA bread. For mandazi, the results suggest that improving farm livelihoods (Item 1), getting energized (Item 15), and providing job opportunities (Item 4) were the easiest items to endorse as a reason for purchase intentions (Table 5). On the other hand, sweetness (Item 18), fiber content (Item 10), crop draught tolerance (Item 8), and nostalgia (Item 14) were the most difficult items to endorse.

Furthermore, in relation to how each item contributes to the functioning of the purchase intention scale as a whole, there was a tendency to under-discriminate for nine items displaying negative values for fit residuals, which contributes to undermining the total score. For seven items, the fit residuals were positive (over-discrimination). Similar to the results for VITA bread, the magnitude of these residuals are not excessive and the findings support the presence of relatively more over-discriminating items at higher levels of item difficulty.

Items 5, 11, 13 and 20 (see Appendix 2 for details) removed because of misfit. SE = standard error. Bonferroni adjustment = 0.003125 (16 items). df = 179.88 for fit residual.

Fig. 1 displays the relationship between the ordinal total score of purchase intentions and their accompanying linear logit measures, together with their logit SEs for each ordinal total score. For both

Table 4
Item difficulty (δ) and fit of the 17 items for VITA bread purchase intentions after rescoring and item elimination, ordered on item difficulty.

Item		Location (difficulty)	SE	Fit residual	Chisq (df = 2)	Prob
4	To create jobs for youths and women	−0.663	0.137	0.010	5.390	0.07
20	It has a high amount of vitamin A	−0.643	0.129	−1.618	6.674	0.04
16	It makes me nostalgic	−0.493	0.124	−0.215	0.373	0.83
11	It prevents the likelihood of eyesight problems	−0.474	0.123	−1.149	5.688	0.06
1	It improves the livelihoods (income) of farm households	−0.350	0.124	0.764	2.977	0.23
2	It improves local production of ingredients	−0.247	0.122	−0.209	1.532	0.46
9	It contains natural colors	−0.190	0.116	0.370	0.987	0.61
18	It has appealing color	0.116	0.110	2.167	1.772	0.41
3	It shortens transportation of inputs for baking	0.142	0.107	−0.985	2.071	0.36
7	It has no added sugar	0.154	0.102	1.651	2.477	0.29
17	It will help my children eat healthy	0.215	0.096	−1.058	0.680	0.71
14	It is tasty	0.236	0.109	0.281	1.059	0.59
10	High fiber content	0.271	0.114	−1.090	2.534	0.28
5	It makes the price of bread affordable	0.391	0.106	0.796	1.686	0.43
12	The package keeps it fresh	0.480	0.110	1.641	2.267	0.32
13	It is soft	0.487	0.097	−0.993	3.160	0.21
8	It promotes growing of a crop that can withstand droughts (climate change)	0.569	0.098	1.518	5.642	0.06

Note: Purchase intentions asked as "I would buy this (OFSP) bread because.".

Table 5
Item difficulty and fit of the 16 items for OFSP mandazi purchase intentions after rescoring and item elimination.

Item		Location (difficulty)	SE	Fit residual	Chisq (df = 2)	Prob
1	Improve the livelihoods (income) of farm households	−1.126	0.128	0.283	0.108	0.95
15	Eating it makes you energetic	−0.777	0.129	0.060	4.141	0.13
4	To create jobs for youths and women	−0.669	0.129	−0.152	1.113	0.57
6	Reduce the vitamin A burden on public health system	−0.558	0.125	−1.024	4.361	0.11
19	It has a high amount of vitamin A	−0.096	0.125	−0.934	3.376	0.18
2	It improves local production of ingredients	−0.094	0.116	−0.641	3.305	0.19
16	It will help my children eat healthy	0.013	0.106	−1.481	5.009	0.08
9	It contains natural colors	0.088	0.106	−0.921	7.046	0.03
7	It has no added sugar	0.198	0.097	−0.296	1.877	0.39
3	Shorten transportation of inputs for baking	0.355	0.102	1.201	1.990	0.37
17	It has appealing color	0.317	0.105	1.048	1.338	0.51
12	It is soft	0.361	0.102	1.225	0.504	0.78
18	Its sweetness	0.495	0.105	1.097	0.707	0.70
10	High fiber content	0.508	0.098	−1.090	1.004	0.61
8	Promote growing of a crop that can withstand droughts (climate change)	0.661	0.094	1.410	5.642	0.08
14	It makes me nostalgic	1.190	0.097	−0.088	3.536	0.17

Note: Purchase intentions asked as "I would buy this (OFSP) bread because…".

products, the results in Fig. 1 show a non-linear relationship between the total score and the transformed logit location. Specifically, the distance between raw scores are wider at the upper and lower ends of the scale of purchase intentions and approach linearity in the middle of the scale range. Furthermore, as Fig. 1 clearly shows, there is no strict ordering of the two distributions of linear logit person locations as they intersect. A Kolmogorov–Smirnov test confirmed that there was no evidence that the locations differed across the two products (test statistic = 0.842, p = 0.477, n = 357).

4.3. *Influence of liking, nutritional beliefs, and socio-demographic variables on purchase intentions*

Fig. 2 shows the conditional inference tree for the relationship between purchase intentions (the transformed Rasch score) and the nine explanatory variables included in the analysis. The liking assessment was the most significant split criterion, splitting the purchase intention measure into two subsets, with higher evaluations on one hand and lower evaluations on the other. For participants with lower evaluation of liking, age was the second most significant splitting criterion, showing that participants older than 28 displayed higher purchase intentions than younger ones. For participants with higher liking evaluations, the purpose of the second split was to further differentiate the rating of purchase intentions between the rating as "like moderately" (2) and "like it very much" (3).

4.4. *Willingness-to-bid and the level of bids*

In this section, we examine whether the Rasch measure of purchase intention helps explain the decisions made in the field auction. Because the decision to bid for an exchange and the decision on how much to bid are chronologically ordered, we also examined whether the decision to bid influenced the actual bid to exchange the products, or whether these decisions are only sequential.

The bidding behavior differed between products. For the VITA bread, 82 participants (51 %) chose to provide a WTP bid, whereas 183 out of 195 participants (93.8 %) did so for the mandazi product. Furthermore, six participants provided bids for their WTA compensation for exchanging breads (average WTA bid = 318 RWF), whereas three participants provided WTA bids for mandazi (average bid = 117). Descriptive statistics for the WTP bids are presented in Appendix 5.

Focusing on the model specification for bids (WTP), we first turn to the test of specification without consideration of the measurement uncertainty in the Rasch score for person location.

For bread, the results suggest that the simpler censored sequential model is more accurate than the two-part hurdle model joint specification (log-likelihood test using the tobit model as the base: 1.69, p = 0.43). Notably, the coefficient for the intercept is significant and strongly negative for bread. With the interpretation going to the uncensored data, this indicates that participants who declined to bid for the VITA bread would have indicated a negative WTP (discount) if it would have been directly possible to do so. For mandazi, the two-part hurdle model joint specification is more accurate than the censored model (log-likelihood test using the tobit model as the base: 6.49, p = 0.039). For both products, the inclusion of the Rasch score improves the fit of the tobit model compared to a model with just an intercept (bread: prob.>chi2 = 0.013; mandazi: prob.>chi2 = 0.003). Moreover, the log(sigma), which is equivalent to the standard deviation of the residual variance, is low compared to the standard deviation of the bid amount (WTP) (Appendix 5: bread = 216.3, mandazi = 118).

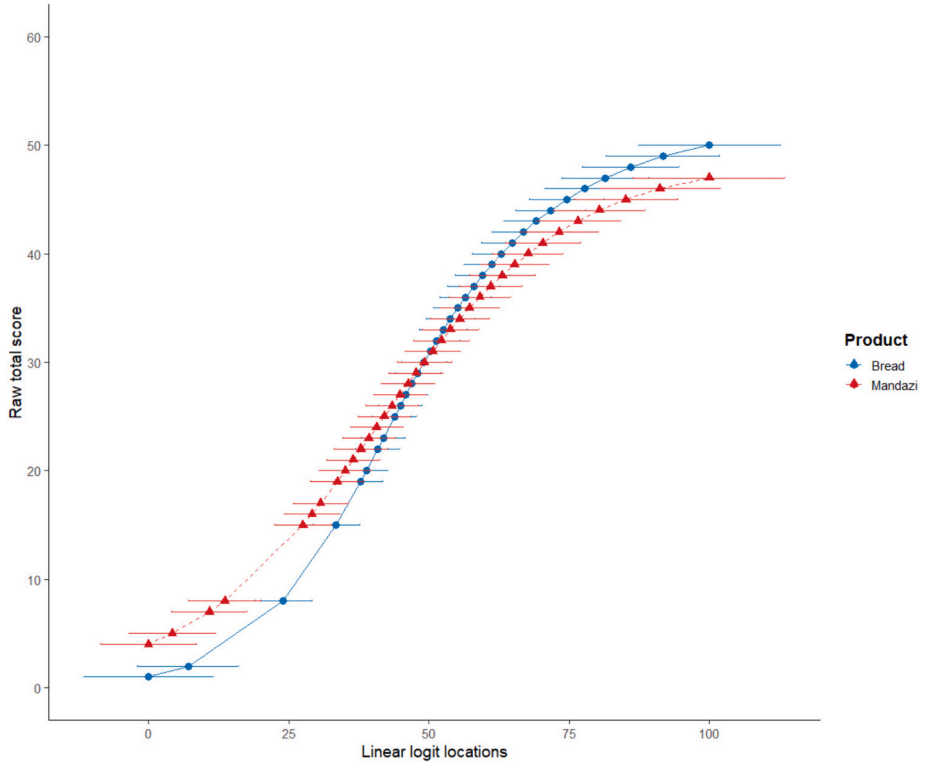


Fig. 1. Relationship between ordinal total score (y-axis) of the purchase intention scale and the corresponding linearized transformed logit locations (x-axis) for VITA bread and OFSP mandazi, respectively. Error bars represent standard error for each logit location.

Based on the findings reported in Fig. 2, we tested the robustness of these findings with the possibility that the effect of the Rasch score on the WTP would also be influenced by age and liking. We included two indicator variables for age > 28 and liking ("like it very much"), as well as an indicator variable for income (>50,000RWF) to control for a potential wealth effect. Interaction variables between each of these indicators and the Rasch score were also included. However, for both bread and mandazi, and based on likelihood tests, the models with only the Rasch score as the explanatory variable were more accurate than the extended models (details available upon request).

We now examine the robustness of our main results by conditioning the regressions of WTP on the Rasch score with the measurement uncertainty as a scale factor using the log of the standard error for the Rasch score (see Eq. 3). Contrary to the results that did not consider measurement uncertainty, the results for bread (see the right-hand part of Table 6a) suggest that the two-part hurdle model is more accurate than the tobit model (log-likelihood test using the conditional tobit model as base: 23.87, $p < 0.001$). Furthermore, and as shown in Table 7, the scaled coefficients indicate that the Rasch score influences the decision to bid and the bid itself differently. This is also confirmed by the AIC and the BIC, which are better for the two-part hurdle model than for the conditional tobit model. Specifically, the effect of the Rasch score, as well as the scale factor on the decision to bid, are stronger than the effect on the level of WTP.

For mandazi, the conditional tobit model is more accurate than the conditional double-hurdle model (6.34, $p = 0.09$, see right-hand part of Table 6b). Furthermore, a comparison based on the AIC and BIC measures supports the argument that the standard (constant zero scale) model provides a better fit to the data than the conditional tobit model, although the differences are not large.

In total, 74 out of 81 participants in the bread arm of the study rated the reasons why they were not willing to provide a bid for exchanging the initially selected bread. For mandazi, this rating was provided by nine out of 12 participants. The main reason for not bidding to exchange was that the products were not considered as popular as other available alternatives within the product category (see Appendix 6).

5. Discussion

Based on an initiative to match with different consumption wants, we examined consumers' preferences for bread and a snack product, both baked using puree from orange-fleshed sweet potato as a partial substitute for flour to obtain vitamin-A-biofortified products. If accepted by consumers, these products can help address the public health issue of vitamin A deficiencies among a vulnerable population in a developing country. Both products are among the most widely consumer breakfast and snack foods and can therefore be effective vehicles/carriers of readily bioavailable vitamin A.

5.1. Multi-item Rasch scales for the measurement of purchase intentions

To the best of our knowledge, the present study is the first to develop and test a Rasch scaling approach to measure food product purchase intentions. This approach makes it possible to evaluate how each product feature characteristic contributes to the purchase intention measure. This becomes of interest for the purpose of identifying key factors for successfully introducing vitamin-A-fortified products to the targeted population and would not be achievable using the classical approach of measuring stated intentions based on a mental readiness to perform the behavior ("I intend to do X") through an evaluative tendency to an object (e.g., Sheeran, 2002).

The first key result of the study is that item difficulties ranged wider for the snack product (mandazi) than for bread, both in terms of strongly deterring purchase intentions and supporting them. Furthermore, for both products, more items qualify as more difficult and fewer items qualify as easy. Despite this, and for both products, the magnitude of difficulties was not excessive. The findings also suggest that the item order differs between the products. In particular, sensory attributes such

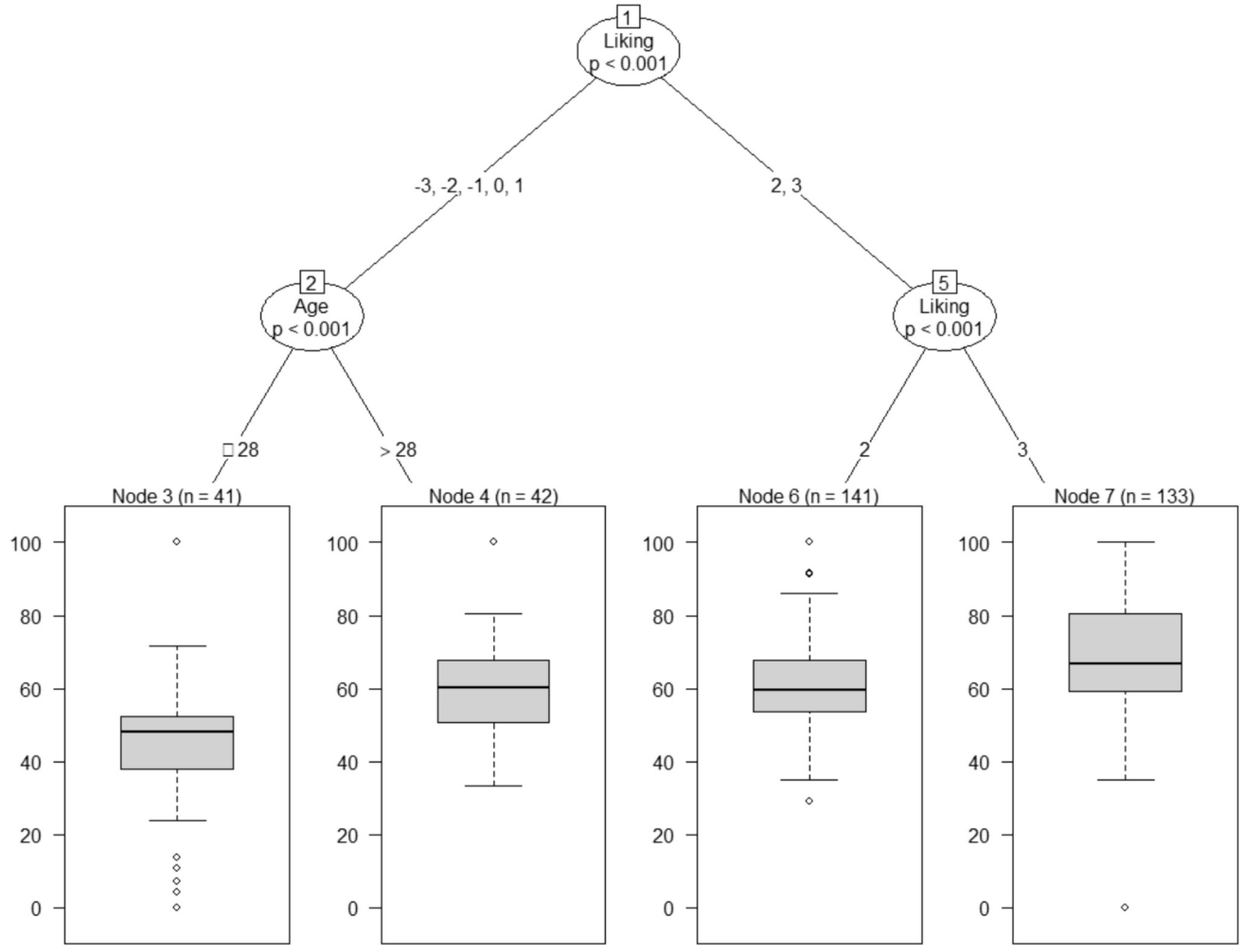


Fig. 2. Conditional inference tree for purchase intentions (transformed Rasch score: min = 0, max =100). For each inner node, input variable and Bonferroni-adjusted *p*-values are given; the bar plot of the Rasch score is displayed for each terminal node. Input variables: product (bread; mandazi), nutritional beliefs [−3:3], actual liking [−3:3], gender, age of respondent, years of education, residence area (urban; peri-urban), income category, and consumption frequency.

as color, softness, sweetness, and fiber content were evaluated as more difficult (that is, expected to inhibit purchases) for mandazi than for bread. Moreover, the attribution of the products to bring a sense of nostalgia had an opposite level of difficulty. In terms of product uniqueness, these findings may serve as input to the further work to align the product formulations to match with the interests of prospective consumers. In doing so, the findings from this study can serve as basis for marketers and the like in terms of which product features to address in communication initiatives.

The Rasch scaling approach also provides measures of person abilities, which then serve as measure of preference intensity. An individual who is located higher in terms of ability is expected to be more likely to be willing to purchase the product. The results suggest that the two product-specific scales yield similar results in terms of the distribution of purchase intentions. Moreover, and importantly for the relationship between purchase intent and economic preferences, the results suggest non-linear relationships in the Rasch scores for person locations and, significantly, that the measurement uncertainty (SE) of person locations is greater for individuals with lower as well as higher total scores on the purchase intention measure. This finding is relevant as an additional explanation in relation to the existing research findings of the discrepancy between stated purchase intentions and WTP.

5.2. *Predicting the Rasch score for purchase intentions*

We adopted a machine learning algorithm to examine the variable selection and the explanatory power of the purchase intention to the ratings of nutritional beliefs, actual liking and a set of socio-demographic variables as covariates, including weekly consumption. Our approach was motivated by a need to examine the stability and sensitivity to persons and subgroups characteristics within the data available, as well as to analyze whether product type, actual liking, and nutritional beliefs served as mediators to the Rasch score measure of purchase intent. Notably, there was no evidence of differentiation in terms of the structural influence of the determinants to purchase intentions between the two products. This finding supports the reliability of the Rasch score measure for purchase intent and suggests that the influence of contextual variables is not decisive to its formation. Moreover, a key result in the present study compared to the existing research on consumer perceptions for vitamin-A-biofortified rice (e.g., Ethen et al., 2024) was that only actual liking and age were significant for the partitioning of purchase intentions. However, when controlling for this result, we found that neither actual liking nor age contributed to being predictive for the decision to bid in the BDM mechanism. In terms of the role of liking, this result fails to corroborate the finding by Kytö et al. (2019). However, the lack of support for nutritional beliefs as a determinant for consumers to choose OFSP bread corroborates the results from Lagerkvist et al.'s (2021) multi-response measurement study.

Table 6
Censored and double-hurdle regressions for WTP in BDM field auction.

a. Bread (n = 162):

		Homoscedasticity			Heteroscedasticity		
		Censored	Two-part hurdle		Censored	Two-part hurdle	
Variable	Definition	Tobit	Logit	Truncated	Tobit	Logit	Truncated
Constant	Intercept	−289.5**	−1.59**	303.7	−290.0**	−97.9	174.7
Rasch score	Purchase intent	5.12**	0.026**	−1.25	5.14**	1.74	1.44
Scale intercept	Scale model				5.29***		5.04***
Rasch SE					−0.01	2.14*	−0.17
logSigma (σ)		5.14***		270.9***			
LL (null)		−651.5					
LL	loglik	−648.4	−109.5	−538.1	−648.4	−106.8	−529.7
df		3	2	3	4	3	4
AIC		1302.8			1304.8†	1286.9†	
BIC		1312.1			1317.2†	1308.6†	

b. Mandazi (n = 195):

		Homoscedasticity			Heteroscedasticity		
		Censored	Two-part hurdle		Censored	Two-part hurdle	
Variable	Definition	Tobit	Logit	Truncated	Tobit	Logit	Truncated
Constant	Intercept	128.8***	1.08	161.6***	124.7***	0.743	136.0***
Rasch score	Purchase intent	1.45***	0.028*	1.11**	1.52***	0.069	1.43***
Scale intercept	Scale model				4.38***		4.51***
Rasch SE					−0.019	0.084	−0.048
logSigma (σ)		4.25***		110.3***			
LL (null)		−1159[c]					
LL	loglik	−1154.4	−43.4	−1107.8	−1154	−42.3	−1109
df		3	2	3	4	3	4
AIC		2313.9			2316.5†	2316.4†	
BIC		2323.7			2329.6†	2339.0†	

*** sign. < 0.01, ** sign. < 0.05, * sign<0.10. SE = Standard error. † AIC and BIC based on re-scaled models (see Section 3.5.3).

Table 7
Scaled coefficients for the censored (Tobit) and double-hurdle conditional regressions for bread.

		Two-part hurdle:	
	Tobit	Probit	Truncated
Intercept	−1.457	−97.86	1.128
Rasch score	0.026	1.745	0.009
Rasch SE	0.015	2.141	−0.173

Note: SE = Standard error.

Regarding the role of age, the finding that younger participants displayed the lowest levels of purchase intentions, conditional upon lower levels of liking, is relevant for the further positioning of the two products to attract consumer interest.

5.3. Combining purchase intentions with choice and economic preferences

The diverging outcome in the first step of the BDM mechanism was rather unexpected, bringing out the potential for product-specific demand differences relevant to the potential success of for market intervention initiatives to promote the adoption of vitamin-A-fortified food products. Therefore, it is relevant to consider how the framing of the BDM mechanism may have contributed to the observed outcome. For example, if bread purchases are more habitual and driven by elements such as affect, trust, brand loyalty, or aspects of product recognition, the use of a self-selected product as the reference point and endowment may have contributed to elicit loss-averse responses, serving as a barrier to contest and inhibiting the decision to be willing to replace the initial bread. In the anticipation of such occurrences, a novel part of the BDM framing included a follow-up part directed at participants not willing to bid for an exchange. This part asked for the interest to instead accept compensation (WTA) to replace products or, alternatively, about the reason for not being willing to replace the self-selected product either in terms of WTP or as WTA. For bread, a majority of the non-bidding participants detailed other reasons to be decisive rather than the possibility of receiving compensation to give up the self-selected bread. This is an interesting finding, which suggests that the issue of loss-aversion would not be resolved by a pricing strategy with discounts for the OFSP bread. However, with little product-specific difference between the reasons for not being willing to replace the self-selected product, it would be reasonable to note that other non-identified aspects relevant to loss aversion may have existed, which were critical to the formation of the reference point and the endowment effects. Further research could seek to identify such reasons. At the same time, it is also relevant to consider how the framing of the BDM mechanism may have contributed to the observed predominance of agreements to bid for the OFSP mandazi product. While this dominance could reflect a halo effect, which would bias the comparative valuation, an alternative, albeit speculative, explanation would go to the snack nature of the mandazi product. Here, issues related to how positive affect to the enjoyable product (category) may have interacted with variety seeking could provide an alternative explanation based on the findings by Kahn and Isen (1993), especially when considering that the price of the existing non-OFSP product used as a reference point was quite low.

Notably, the product discrepancy in terms of decisions to bid in the BDM contrasts with the non-separation outcome for purchase intentions, which was noted between the two products using the Rasch scale measure. This product comparison contributes to existing field auction research (Barber et al., 2012) and challenges the notion of product-based consistency between the stated preference measure based on an entity between attitudes and behaviors, on one hand, and economically consequential behavior, on the other. Relevant to potential pricing strategies for the further introduction of OFSP biofortified foods, these results suggest risks of so-called preference reversals. As Kahneman et al. (1999) noted, the risks of reversals ought to be higher when two

conceptually equivalent methods for eliciting product preference are conflicting.

Next, a key result of the present study is that consideration of the measurement uncertainty from the Rasch scale for purchase intentions has important implications for the relationship between purchase intentions and the decisions made in the BDM. The key question here is whether the purchase intent explains the two decisions within the BDM in the same way, as well as whether the amount to bid is conditional upon (influenced by) the decision to bid.

Without considering the measurement uncertainty, for mandazi and given the predominance for participants to be willing to bid for this product in the BDM, there is support for the argument that the choice to bid influenced the decision on the amount to bid. Furthermore, the effects of the Rasch score were positive for both decisions in the BDM, but weaker on the choice to bid and instead more influential on the amount of WTP. This influence of the bid response in the real context suggests a highly positive (inflated) disposition for the biofortified mandazi product. Therefore, it remains an open question as to whether consumers with intentions to buy the OFSP mandazi would be prepared to pay a mark-up to do so, or if this tendency to exaggerate the WTP in the BDM mechanism is due to a halo effect. For bread, and given the outcome of decisions to bid for the exchange, the appropriateness of the hurdle model would be reasonable considering the clear divide between participants in terms of choice to bid to exchange. The hurdle model would focus only on the subset of participants who signaled their preferences for the product in the real context. However, there was no evidence supporting separate effects for the Rasch score on the two decisions, as the hurdle model provided a less accurate fit to data than the censored model. Hence, for the OFSP bread, the decision to bid and the level of the bid provided are driven by the same latent process and the decision about the amount to bid is not influenced by the decision to bid.

Notably, when the measurement uncertainty for person locations in the Rasch scale is considered, and as a novel contribution, we found that the two-part hurdle model is more accurate than the tobit model for bread. Although the Rasch score for purchase intentions was not statistically significant in the estimation, the results (with the interpretation based on scaled regression coefficients) indicate that purchase intentions drive the choice to bid and the amount to bid differently. The score on the Rasch scale, and the variability in this measure, has a stronger effect on the choice to bid, but a substantially weaker effect on the bid amount. Furthermore, for mandazi, the results suggest that the unconditional tobit model was more appropriate. Hence, these findings suggest that it is more important to consider measurement uncertainty in purchase intentions when the comparative evaluation (first part of the BDM) that it helps explain is more widely dispersed across the study sample, and less influential when the outcome is more coherent.

6. Conclusion

A central query for research analyzing food consumer preferences is whether the applied methods serve to predict actual choices within the in-store shopping context. Therefore, with a multi-response approach to product research, the present research developed and tested a measure for purchase intentions based on Campbell's paradigm (Campbell, 1963; Kaiser et al., 2010). This measure details how product item characteristics and person abilities contribute to the formation of purchase intentions. Moreover, we investigated the extent to which the purchase intent influences the decision made within an in-store auction to elicit monetary preferences in terms of product exchange, and also whether the monetary preference in such a case is influenced by the decision to be willing to exchange. We also investigated the structural relationship between purchase intentions and multiple measures of potential drivers for purchase intentions.

The key finding of this study is that the nature of the relationship between the BDM approach and the stated purchase intention measure is product-specific because of differences in the outcome of the BDM. The nature of the relationship also depends on whether or not the measurement uncertainty in the purchase intentions score for person locations is considered. In this study, inclusion of the measurement uncertainty reverses the extent to which the choice to bid and the bid amount within the BDM are driven by separate processes. The primary insight is that consideration of the measurement uncertainty in the purchase intention measure is relevant when the auction outcome is mixed (for example, the almost 50/50 mix in willingness to bid observed for the vitamin-A-biofortified bread). In this situation, there was no evidence that purchase intensions influenced the decisions within the BDM. On the other hand, the measurement uncertainty for purchase intentions holds no informational value when the auction outcome gives strong dominance for one product over the other (such as for the OFSP Mandazi). In this situation, the results support the argument that stated and actual preferences are both indicators of the same latent disposition, although the effect size of purchase intention is modest.

The multi-response approach adopted in this study supports the argument that the sensory attributes for the vitamin-A-biofortified products do not serve as key product features to promote (or detract) purchase intentions. Addressing and aligning sensory and product characteristics should be a priority in further developing product formulations and marketing of vitamin-A-biofortified products to better match with consumer motivations. Moreover, while actual liking helped predict purchase intentions, we found no evidence of such an association for nutritional beliefs (after information about vitamin A enrichment and after tasting). This intriguing finding suggests that population nutrition education and awareness creation, and sensitization regarding the benefits of vitamin-A-biofortified products as a source of essential nutrients and not just energy, should be a priority in order to build congruent support when attempting to target vulnerable households by matching the nutrient enrichment to in-demand products.

CRediT authorship contribution statement

Carl-Johan Lagerkvist: Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Julius J. Okello:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization. **Sylvester Ojwang:** Writing – review & editing, Supervision, Investigation, Data curation. **Herve Mwizerwa:** Supervision, Investigation. **Jean Claude Nshimiyimana:** Writing – review & editing, Supervision, Investigation. **Jean Ndirigwe:** Writing – review & editing, Supervision, Investigation. **John Bosco Shingiro:** Writing – review & editing, Supervision, Investigation. **Jan Low:** Writing – review & editing, Funding acquisition.

Ethical statement

This study was conducted according to the guidelines set out in the Declaration of Helsinki and all procedures involving human subjects were approved by the National Institute of Statistics in Rwanda on Dec 2, 2022 (No. 0532/2022/10/INSR).

Funding statement

Financial support for this study was received from West and Central Africa CGIAR Regional Initiative. Additional funding for the study was received from the CGIAR Initiative on Market Intelligence.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was implemented in partnership with the Rwanda Agriculture and Animal Resource Development Board (RAB), and orange-fleshed-sweetpotato (OFSP) puree product processors; namely, URWIBUTSO Enterprises, which processes OFSP-based mandazi, and the CARL Group, which processes the VITA bread. These organizations and businesses have not had any influence on the results that are presented in this manuscript. We wish to thank four anonymous reviewers for substantive comments on earlier versions of this article.

Appendix A. Appendix 1

A.1. Product Narrative

This [OFSP] bread/mandazi has been made from sweetpotato and wheat flour. The sweetpotato used is produced locally by smallholder farmers and is rich in vitamin A, an essential nutrient that is crucial for maternal health and child survival. Vitamin A deficiency (VAD) can lean to severe visual impairment and blindness, and significantly increases the risk of severe illnesses, and even death, from such common childhood infections as diarrhea, and measles among children.

Vitamin A deficiency (VAD) remains a major public health problem in Rwanda and many other African countries. One of the immediate causes of VAD is inadequate dietary intake of foods that are rich in vitamin A, such as the orange-fleshed sweetpotato (OFSP), by the vulnerable groups. Thousands of preschool children and pregnant women are currently at risk of VAD in Rwanda. Pregnant women are more vulnerable to VAD during the last three months of their pregnancy, when demand by both the unborn child and the mother is highest.

This [OFSP] bread/mandazi offers the opportunity to fight VAD. It is made by a bakery/company called 'CARL Group' based in Kigali.

Appendix B. Appendix 2

Table A.2
Items and responses per category for purchase intentions.

Item		Bread ($n = 164$)					Mandazi ($n = 218$)				
		1	2	3	4	5	1	2	3	4	5
1	Growing OFSP as an input into [product] making will improve the livelihoods of farm households	1	5	8	56	94	0	6	19	58	135
2	Local production of […] will be improved	2	3	21	74	64	1	5	37	94	81
3	Shorten transportation of inputs for baking […]	5	6	20	64	69	8	8	32	74	96
4	Encourage sweet potato farming to create jobs for youth and women	2	2	6	52	102	2	4	11	64	137
5	Promote usage of sweet potato in […] to make the price affordable	3	11	18	77	55	7	15	34	81	81
6	Consumption of VITA bread/OFSP mandazi helps reduce the vitamin A burden on public health system	2	8	10	54	90	0	4	13	61	140
7	It has no added sugar	9	4	13	56	82	5	15	31	38	129
8	It will promote growing of a crop that can withstand droughts (climate change)	4	13	38	60	49	10	9	65	57	77
9	It contains natural colors	2	3	22	65	72	2	10	25	71	110
10	It has high fiber content	4	6	25	80	49	4	12	48	67	87
11	Eating it regularly reduces the likelihood of eyesight problems	2	4	10	37	111	3	5	22	46	142
12	The package keeps it fresh	10	3	22	87	42	N. A.	N. A.	N. A.	N. A.	N. A.
13	It is soft and is therefore appreciated by my children	16	3	23	61	61	23	1	27	78	89
14	It is tasty	4	7	23	71	59	1	12	30	88	87
15	It makes me nostalgic	15	15	38	64	32	8	20	62	76	52
16	Eating it makes me energetic	3	1	15	60	85	0	3	14	49	152
17	It will help my children eat healthy	15	1	18	43	87	19	1	20	56	122
18	It has appealing color	4	5	21	72	62	5	7	32	87	87
19	Of its sweetness	10	19	29	61	45	2	12	28	97	79
20	It has a high amount of vitamin A	2	3	10	35	114	0	3	18	43	154

Note: Response categories: (1) I completely disagree, (2) I somewhat disagree, (3) I neither agree or disagree, (4) I somewhat agree, (5) I fully agree. Statements in relation to: "I intend to buy this [VITA bread/ OFSP mandazi] because/to". For mandazi, Item 12 was not included as this product is not packaged. Number of respondents (n) based on complete samples.

Appendix C. Appendix 3

C.1. Instructions for BDM field auction with bid to exchange

Before this interview, you selected to buy [name of bread/mandazi] (henceforth Bread1/Mandazi1) and during this interview, you have tasted the VITA BREAD/OFSP Mandazi.

As a token of appreciation for your participation in this study, we are giving you [bread = 2000 RWF; mandazi = 500RWF].

There is now a chance for you to exchange (Bread1/Mandazi1) and instead leave this place with the [VITA BREAD/OFSP Mandazi]. Doing so will require you to be willing to use some of the money you have just received to purchase the [VITA BREAD/OFSP Mandazi]. A random draw will be made to determine the extra amount for which the exchange will take place. You will leave (Bread1/Mandazi1) with us if the exchange takes place. We will provide details and examples in a moment in case you are interested in an opportunity to exchange (Bread1/Mandazi1) for the [VITA Bread/OFSP Mandazi].

Are you interested in this option?

1 = Yes, I would like to have the opportunity to get the VITA Bread/OFSP Mandazi replacing (Bread1/Mandazi1).

0 = No, I'll keep Bread 1/Mandazi 1 → Proceed to next page.

If "Yes":

Now, please think of how much of the 2000 RWF (500 RWF *in the case of mandazi*) you are willing to pay right now to cover the amount needed to replace [Bread1/Mandazi1] with VITA BREAD (OFSP Mandazi). Once you decide the amount, we will make a random draw of numbers between **0 and 2000** (between **1 and 500** *in the case of mandazi*). You do the exchange for the VITA BREAD (OFSP Mandazi) if the number we draw is lower than or equal to the amount that you stated. Therefore, please be sure that you indicate the amount of money that, according to you, matches the difference in price you are willing to pay between the two types of bread.

An example in the case of bread: If you indicate that you are willing to pay 600 RWF to replace Bread1, and our draw gives 400 RWF ≥ you will pay the 600 RWF and take the VITA BREAD home +1400 RWF – the unused token of appreciation of your time in the study. But … if our draw had been from 700 RWF or above, you would instead keep Bread 1 + 2000 RWF.

An example in the case of Mandazi: If you indicate that you are willing to willing to pay 300 RWF to replace Mandazi 1, and our draw gives 200 RWF ≥ you will pay the 300 RWF and take the OFSP Mandazi home +200 RWF – the unused token of appreciation of your time in the study. But… if our draw had been from 400 RWF or above, you would instead keep Mandazi1 + the 500 RWF token of appreciation.

Please, now give the amount that you are willing to pay to replace [Bread1/Mandazi1]:

Amount willing to pay: RWF_____.

If "No":

We understand that you are not willing to pay any extra money to have the chance to get the [VITA bread/OFSP Mandazi].

Would you be interested in agreeing to leave here with [VITA bread/OFSP Mandazi] for a certain compensation? Think of this as a discount that you would seek in order to exchange the [bread/mandazi] you initially selected.

1 = Yes; 0 = No.

If "Yes".

[***In the case of bread***] : Please now give an amount between 1 and 360 RWF that you are willing to accept as compensation to replace breads. Please note that once you provide the price/amount, we will randomly draw a number between 1 and 360. If the number we draw is LESS than the amount of money you stated, you take the VITA bread home and get paid the amount you stated + the 2000RWF that were provided as token of appreciation for your participation in this study. If the number we draw is MORE than the amount of money you stated, you will go home with your Bread 1 and the token of appreciation (2000 RWF) without any additional compensation.

[***In the case of mandazi***]: Please now give the amount between 1 and 150 RWF that you are willing to accept as compensation to replace mandazis. Please note that once you provide the price/amount, we will randomly draw a number between 1 and 150. If the number we draw is less than or equal to the amount of money you stated, you take the OFSP mandazi home and get paid the amount you stated + the 500 that were provided as token of appreciation for your participation in this study. If the number we draw is more than the amount of money you stated you lose the OFSP Mandazi, go home with your Mandazi 1 and the token of appreciation (500 RWF) without any additional compensations.

If "No":

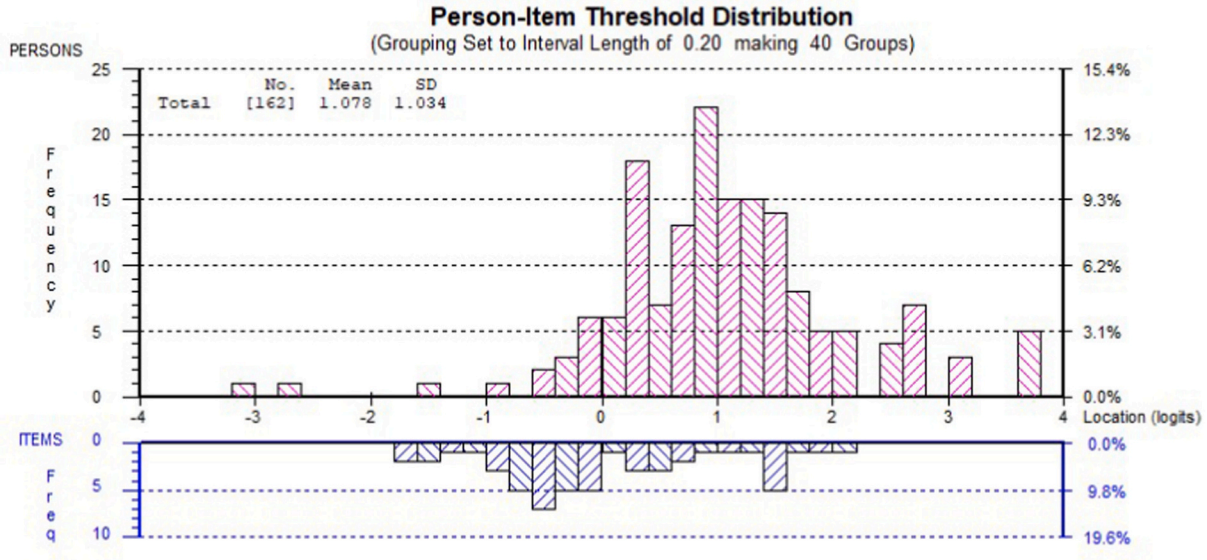
Please indicate the extent to which you agree or disagree with the following as the main reasons for you not to be willing to replace your initially selected bread(mandazi) with the VITA bread (OFSP Mandazi) using the scale below:

1 = Strongly Disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Strongly Agree.

	1	2	3	4	5
a) I do not like sweet potatoes					
b) I do not eat sweet potatoes					
c) I am not aware of the importance of vitamin A					
d) I doubt vitamin A is that important					
e) This bread/mandazi is not as popular as other breads/mandazis					
f) I get heartburn and/or stomach upset when I eat sweet potatoes					
g) I do not care about the nutritional content of the bread/mandazi					
h) The bread/mandazi is not affordable to me					
i) I do not like the sweet taste of the bread/mandazi. I prefer a salty taste.					

Appendix D. Appendix 4

(a) Bread:



(b) Mandazi

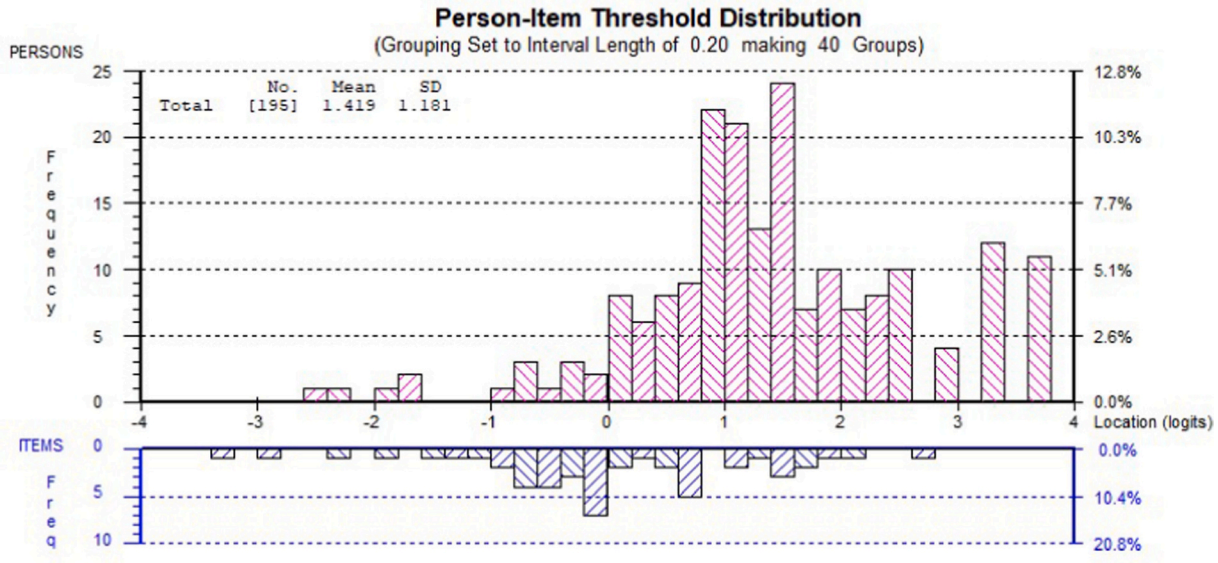
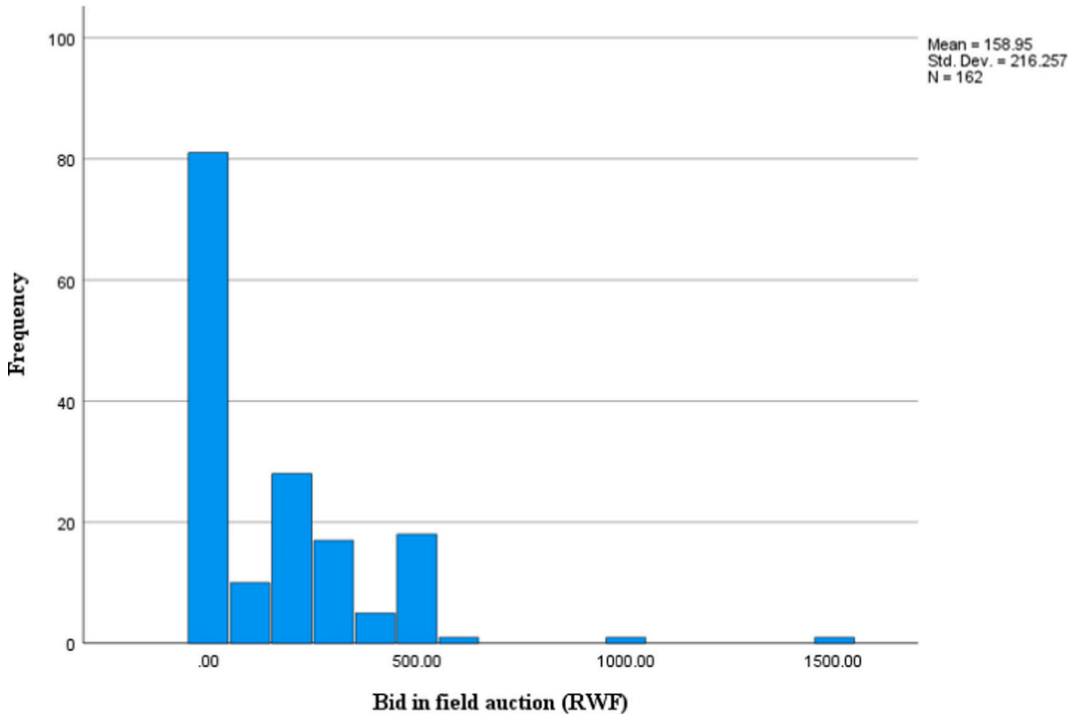


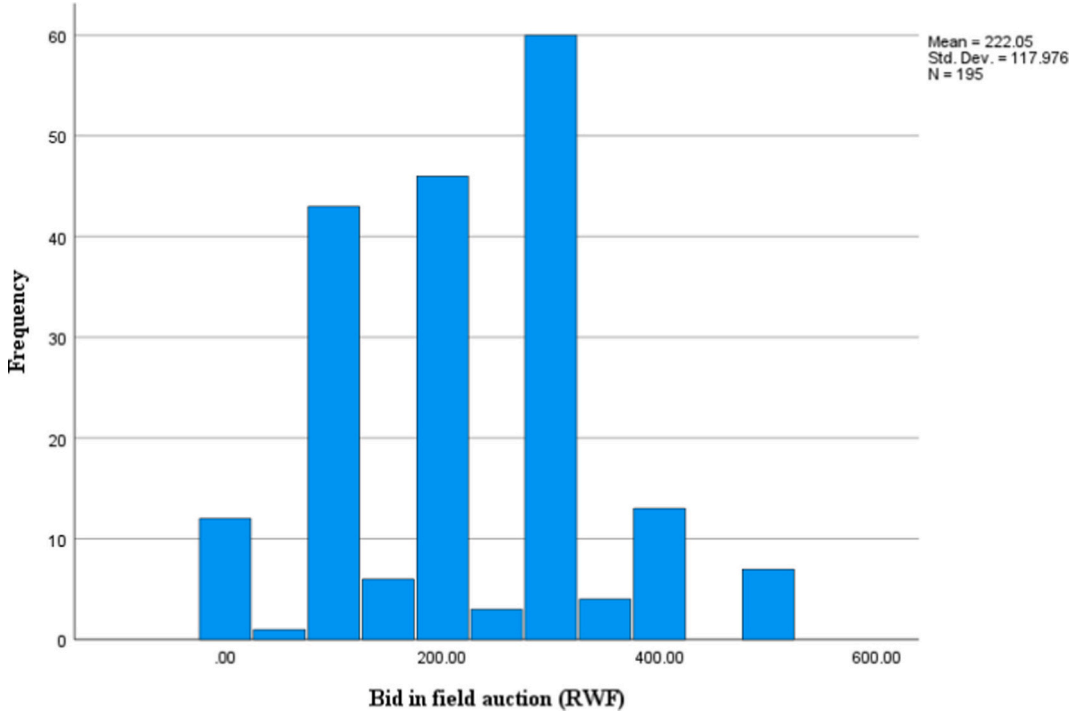
Fig. A.3.1. Person-Item Rasch threshold distributions for (a) Bread and (b) Mandazi.

Appendix E. Appendix 5



Note: Available bid range [0, 2,000]

Fig. A.4.1. Frequency of bids to replace initially selected bread for VITA bread.
Note: Available bid range [0,2000].



Note: Available bid range [0, 500]

Fig. A.4.2. Frequency bids to replace initially selected mandazi for OFSP Mandazi.
Note: Available bid range [0, 500].

Appendix F. Appendix 6

Please indicate the extent to which you agree or disagree with the following as the main reasons why you would not be willing to replace your initially selected bread (mandazi) with the VITA bread (OFSP Mandazi):

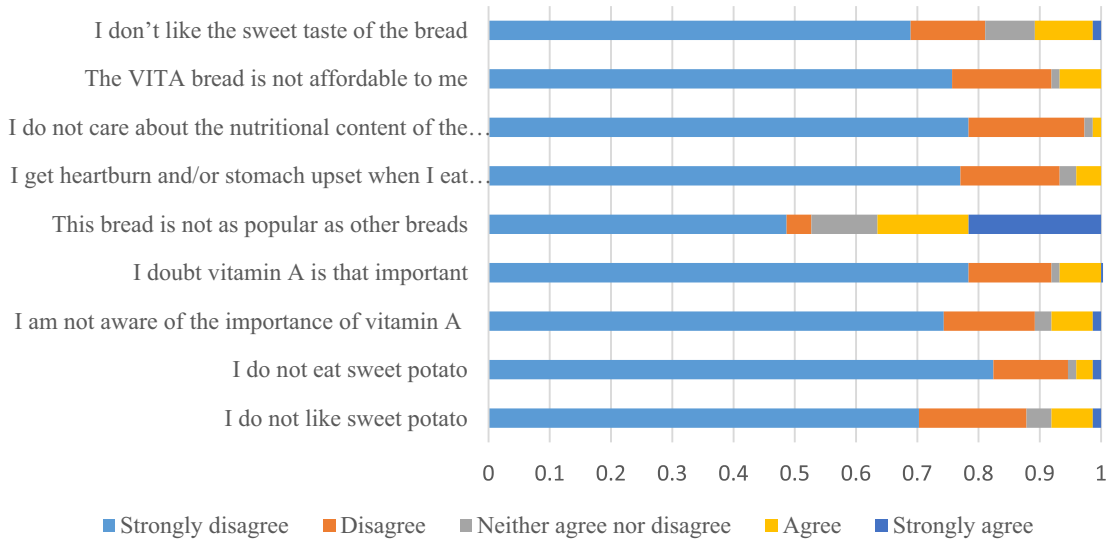


Fig. A.6.1. Agreement with reasons for declining to bid to replace the initially selected bread with VITA bread in the Becker-DeGroot-Marchak auction ($n = 74$: 45.7 % out of 162).

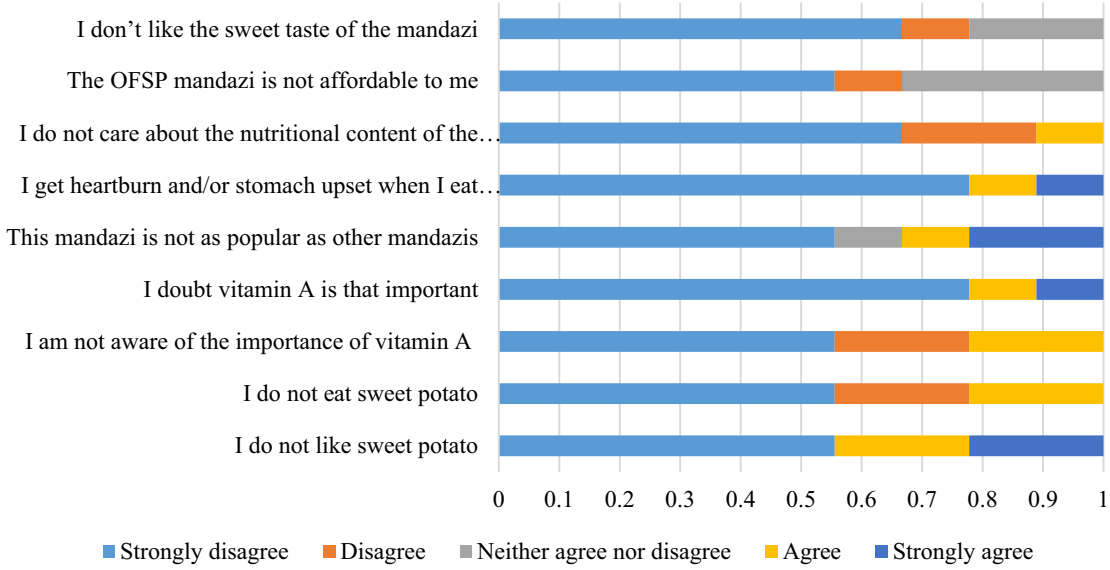


Fig. A.6.2. Agreement with reasons for declining to bid to replace the initially selected non-OFSP mandazi with OFSP mandazi in the Becker-DeGroot-Marchak auction ($n = 9$: 4.6 % out of 195).

Data availability

Data will be made available on request.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior & Human Decision Processes, 50*, 179–211.

Ajzen, I., Brown, T. C., & Carvajal, F. (2004). Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation. *Personality and Social Psychology Bulletin, 30*(9), 1108–1121.

Ajzen, I., & Driver, B. L. (1992). Contingent value measurement: On the nature and meaning of willingness to pay. *Journal of Consumer Psychology, 1*, 297–316.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561–573. https://doi.org/10.1007/bf02293814

Andrich, D., Lyne, A., Sheridan, B., & Luo, G. (2010). *RUMM2030Plus*. Perth: RUMM Laboratory.

Armstrong, J. S., Morwitz, V. G., & Kumar, V. (2000). Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy? *International Journal of Forecasting, 16*, 383–397.

Ballco, P., & Gracia, A. (2022). Tackling nutritional and health claims to disentangle their effects on consumer food choices and behavior. A systematic review. *Food Quality and Preference, 101*, Article 104634.

Barber, N., Kuo, P.-J., Bishop, M., & Goodman, R., Jr. (2012). Measuring psychographics to assess purchase intention and willingness to pay. *Journal of Consumer Marketing, 29*(4), 280–292.

Becker, G. M., De-Groot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioural Science, 9*, 226–232.

Bennett, J., & Blamey, R. (Eds.). (2001). *The choice modelling approach to environmental valuation*. Edward Elgar Publishing.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth.

Campbell, D. T. (1963). Social attitudes and other acquired behavioral dispositions. In S. Koch (Ed.), *Vol. 6. Psychology: A study of a science* (pp. 94–172). New York, NY: McGraw-Hill.

Caputo, V., Lagoudakis, A., Shupp, R., & Bazzani, C. (2023). Comparing experimental auctions and real choice experiments in food choice: A homegrown and induced value analysis. *European Review of Agricultural Economics, 50*(5), 1796–1823.
Carrol, K. A., & Samek, A. (2018). Field experiments on food choice in grocery stores: A 'how-to' guide. *Food Policy, 79*, 331–340.
Croissant, Y., & Zeileis, A. (2018). Truncreg: Truncated gaussian regression models. *R package version 0*, 2–5. https://r-forge.r-project.org/projects/truncreg.
Dalenberg, J. R., Gutjar, S., ter Horst, G. J., de Graaf, K., Renken, R. J., & Jager, G. (2014). Evoked emotions predict food choice. *PLoS One, 9*, Article e115388.
De Groote, H., Chege, C. K., Tomlins, K., & Gunaratna, N. S. (2014). Combining experimental auctions with modified home-use test to assess rural consumers' acceptance of quality protein maize, a biofortified crop. *Food Quality and Preference, 38*, 1–13.
De Groote, H., Kariuki, S. W., Traore, D., Taylor, J. R. N., Ferruzzi, M. G., & Hamaker, B. R. (2018). Measuring consumers' interest in instant fortified pearl millet products: A field experiment in Touba, Senegal. *Journal of the Science of Food and Agriculture, 98*, 2320–2331.
DeFleur, M. L., & Westie, F. R. (1958). Verbal attitudes and overt acts: An experiment on the salience of attitudes. *American Sociological Review, 23*, 667–673.
Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
Ekstrand, J., Westergren, A., Årestedt, K., Hellström, A., & Hagell, P. (2022). Transformation of Rasch model logits for enhanced interpretability. *BMC Medical Research Methodology, 22*, 332. https://doi.org/10.1186/s12874-022-01816-1
Ethen, D. Z., Alam, M. J., & De Steur, H. (2024). Unfolding consumer purchase intentions toward genetically modified rice with nutritional benefits in Bangladesh. *Outlook on Agriculture, 53*(3), 277–290.
Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (2008). On the automatic activation of attitudes. In R. H. Fazio, & R. E. Petty (Eds.), *Attitudes: Their structure, function, and consequences* (pp. 17–32). New York, NY: Psychology Press.
Fischer, A. R. H., & Frewer, L. J. (2009). Reliability of the Rasch food safety practices scale. *Appetite, 53*, 241–244.
Fischer, A. R. H., Frewer, L. J., & Nauta, M. N. (2006). Toward improving food safety in the domestic environment: A multi-item Rasch scale for the measurement of the safety efficacy of domestic food-handling practices. *Risk Analysis, 26*(5), 1323–1338.
Forde, C. G., & de Graaf, K. C. (2023). Sensory influences on food choice and energy intake: Recent developments and future directions. In E. Guichard, & C. Salles (Eds.), *Flavor* (2nd ed., pp. 329–362). Woodhead Publishing. https://doi.org/10.1016/B978-0-323-89903-1.00013-X.
Giacalone, D., & Jaeger, S. R. (2021). Sensory drivers of perceived situational appropriateness in unbranded foods and beverages: Towards a deeper understanding. *Appetite, 167*, Article 105589.
Hagell, P. (2014). Testing rating scale unidimensionality using the principal component analysis (PCA)/*t*-test protocol with the Rasch model: The primacy of theory over statistics. *Open Journal of Statistics, 4*, 456–465.
Hagquist, C., Bruce, M., & Gustavsson, J. P. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies, 46*, 380–393.
Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics, 15*(3), 651–674.
Hothorn, T., & Zeileis, A. (2015). Partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research, 16*, 3905–3909.
Irwin, J. R., McClelland, G. H., McKee, M., Schulze, W. D., & Norden, N. E. (1998). Payoff dominance vs. cognitive transparency in decision making. *Economic Inquiry, 36*, 272–285.
Jamieson, L. F., & Bass, F. M. (1989). Adjusting stated intention measures to predict trial purchase of new products: A comparison of models and methods. *Journal of Marketing Research, XXVI*, 336–345.
Kahn, B. E., & Isen, A. M. (1993). The influence of positive affect on variety seeking among safe, enjoyable products. *Journal of Consumer Research, 20*(2), 257–270.
Kahneman, D., & Ritov, I. (1994). Determinants of stated willingness to pay for public goods: A study in the headline method. *Journal of Risk and Uncertainty, 9*, 5–38.
Kahneman, D., Ritov, I., & Schkade, D. (1999). Economic preferences or attitude expressions?: An analysis of dollar responses to public issues. *Journal of Risk and Uncertainty, 19*(1), 203–235.
Kaiser, F. G., Byrka, K., & Hartig, T. (2010). Reviving Campell's paradigm for attitude research. *Personality and Social Psychology Review, 14*(4), 351–367.
Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics, 29*, 119–127.
Kytö, E., Virtanen, M., & Mustonen, S. (2019). From intention to action: Predicting purchase behavior with consumers' product expectations and perceptions, and their individual properties. *Food Quality and Preference, 75*, 1–9.
Lagerkvist, C. J., Mwende, J., Muoki, P., & Okello, J. J. (2021). How useful are perception- and experienced-based measures in predicting actual food choice? Evidence from an in-store field experiment using a multi-response approach. *Food Quality and Preference, 94*, Article 104320.
Lagerkvist, C. J., & Okello, J. J. (2016). Using the integrative model of behavioral prediction and censored quantile regression to explain consumers' revealed preferences for food safety: Evidence from a field experiment in Kenya. *Food Quality and Preference, 49*, 75–86.
Logue, A. W. (1986). *The psychology of eating and drinking*. New York: Freeman.
Low, J., & Van Jaarsveld, P. (2008). The potential contribution of bread buns fortified with β-carotene-rich sweet potato in Central Mozambique. *Food and Nutrition Bulletin, 29*(2), 98–107.
Naico, A. T. A., & Lusk, J. L. (2010). The value of a nutritionally enhanced staple crop: Results from a choice experiment conducted with orange-fleshed sweet potatoes in Mozambique. *Journal of African Economies, 19*(4), 536–558.
Lusk, J. L., & Shogren, J. F. (2007). *Experimental auctions: Methods and applications in economic and marketing research*. Cambridge University Press.
Marzilli Ericson, K. M., & Fuster, A. (2011). Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *The Quarterly Journal of Economics*, 1–29.
Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149–174.
Messner, J. W., Mayr, G. J., & Zeileis, A. (2016). Heteroscedastic censored and truncated regression with crch. *The R Journal, 8*(1), 173–181.
Muoki, P., & Kwikiriza, N. (2019). *Incentives and disincentives to investment in orange-fleshed sweetpotato – The case of a pro-poor public–private partnership in Rwanda. Technical assessment of orange-fleshed sweetpotato value chain in Rwanda. Scaling up Sweetpotato through agriculture and nutrition project (SUSTAIN). International potato Center* (p. 34). https://doi.org/10.4160/9789290605133
Mustonen, S., Hissa, I., Huotilainen, A., Miettinen, S., & Tuorila, H. (2007). Hedonic responses as predictors of food choice: Flexibility and self-prediction. *Appetite, 49*, 159–168.
Muzhingi, T., Langyintuo, A. S., Malaba, L. C., & Banzinger, M. (2008). Consumer acceptability of yellow maize in Zimbabwe. *Food Policy, 33*, 352–361.
Okello, J. J., Shiundu, F. M., Mwende, J., Lagerkvist, C. J., Nyikal, R. A., Muoki, P., & Heck, S. (2021). Quality and psychosocial factors influencing purchase of orange-fleshed sweetpotato bread. *International Journal of Food Science & Technology, 56*(3), 1432–1446.
Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the hospital anxiety and depression scale (HADS). *British Journal of Clinical Psychology, 46*, 1–18.
Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: The University of Chicago Press.
Seibold, H., Zeileis, A., & Hothorn, T. (2016). Model-based recursive partitioning for subgroup analysis. *International Journal of Biostatistics, 12*(1), 45–63.
Shogren, J. F., Cho, S., Koo, C., List, J., Park, C., Polo, P., & Wilhelmi, R. (2001). Auction mechanisms and the measurement of WTP and WTA. *Resource and Energy Economics, 23*(2), 97–109.
Smith, E. V., Jr. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement, 3*, 205–231.
Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis and Rheumatology, 57*, 1358–1362.
Therneau, T. M. (2020). A package for survival analysis in R. https://CRAN.R-project.org/package=survival.
Vecchio, R., & Borrello, M. (2019). Measuring food preferences through experimental auctions: A review. *Food Research International, 116*, 1113–1120.
Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.
World Health Organization (WHO). (2019). Micronutrient deficiencies. Vitamin A deficiency. https://apps.who.int/iris/bitstream/handle/10665/44664/9789241501767_eng.pdf?sequence=1 accessed 10/10/2024.
Xue, H., Mainville, D., You, W., & Nayga, R. M., Jr. (2010). Consumer preferences and willingness to pay for grass-fed beef: Empirical evidence from in-store experiments. *Food Quality and Preference, 21*, 857–866.
Zeileis, A., Koenker, R., & Doebler, P. (2024). Glmx: Generalized linear models extended. *R package version 0.2–1*. https://CRAN.R-project.org/package=glmx.
Zeileis, A., Messner, J. W., Stauffer, R., Kosmidis, I., & Mayr, G. J. (2024). *Crch: Censored regression with conditional heteroscedasticity, 1.2–1*. https://doi.org/10.32614/CRAN.package.crch
Zhao, T., Liu, S., Zhang, R., Zhao, Z., Pu, L., Wang, L., & Han, L. (2022). Global burden of vitamin a deficiency in 204 countries and territories from 1990–2019. *Nutrients, 14*, 950. https://doi.org/10.3390/nu14050950