RESEARCH ARTICLE

Weighting or aggregating? Investigating information processing in multi-attribute choices

Mesfin G. Genie[1,3] | **Nicolas Krucien[2]** | **Mandy Ryan[1]**

[1]Health Economics Research Unit, University of Aberdeen, Aberdeen, UK

[2]Patient-Centered Research, Evidera, London, UK

[3]Department of Economics, Ca' Foscari University of Venice, Venice, Italy

Correspondence
Mesfin G. Genie, Health Economics Research Unit, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, AB25 2ZD.
Email: mesfin.genie@abdn.ac.uk

Abstract

Multi-attribute choices are commonly analyzed in economics to value goods and services. Analysis assumes individuals consider all attributes, making trade-offs between them. Such decision-making is cognitively demanding, often triggering alternative decision rules. We develop a new model where individuals aggregate multi-attribute information into meta-attributes. Applying our model to a choice experiment (CE) dataset, accounting for attribute aggregation (AA) improves model fit. The probability of adopting AA is greater for: homogenous attribute information; participants who had shorter response time and failed the dominance test; and for later located choices. Accounting for AA has implications for welfare estimates. Our results underline the importance of accounting for information processing rules when modelling multi-attribute choices.

KEYWORDS
attributes aggregation, choice experiment, choice modelling, information processing, multi-attribute choices

JEL CLASSIFICATION
C35, D01, D80, D90

1 | INTRODUCTION

Based on Samuelson's theory of revealed preferences (Samuelson, 1938), both actual and hypothetical choices are frequently used in applied economics to measure sensitivities of the demand to marginal changes in goods and services. When actual choices (also referred as revealed preferences) are either unavailable (e.g., no market data for an innovation) or imperfect (e.g., patients' choices are often driven by medical recommendations and/or regulatory limitations), nonmarket valuation techniques are used to observe hypothetical consumption decisions (also referred as stated preferences) (Ben-Akiva et al., 1985; Louviere et al., 2000). Both actual and hypothetical choice data are analyzed within the random utility maximization (RUM) framework to obtain marginal utilities (McFadden, 1974).

At the core of RUM is the Lancasterian theory of demand (LTD) (Lancaster, 1966), stipulating that individuals consider the features (or attributes) of the good when choosing between alternatives, rather than the good *per se*. The utility an individual derives from a particular good, therefore, depends on the relative importance of its features.

However, the initial formulation of the LTD does not stipulate how individuals combine these different pieces of information. Studies typically assume that individuals follow an attribute weighting (AW) rule when processing multi-attribute information. The utility of a good is a weighted average of its features with weights defined by the individuals' preferences. AW is popular, ensuring a continuous (and differentiable) utility function that allows for *easy* computation of marginal rates of substitution (MRS) and choice elasticities (Araña & León, 2009; Borghans et al., 2010; Breustedt et al., 2008; de Bekker-Grob et al., 2012; Hackbarth & Madlener, 2013; Hansen, 1976; Hensher, 1994; Hoyos, 2010; Lancsar & Louviere, 2008; Ryan et al., 2007; Scott, 2002).

However, the behavioral validity of the AW rule can be challenged. AW implies individuals can make a relatively large number of trade-offs between the different features of the good or service being valued. For instance, when K features are used to describe the choice options, and assuming that individuals would consider all these features, then a maximum of $K \times (K-1)/2$ trade-offs exist. With each trade-off being cognitively demanding (Luce et al., 1999), individuals are likely to use *simplifying* decision rules (Simon, 1957). For example, applying a satisficing rule, individuals process the multi-attribute information of each option one at a time and stop when they identify a "*good enough*" option. Satisficing allows individuals to process only a subset of the available information. Alternatively, rather than modifying the amount of information to be processed, other rules simplify the decision-making by alleviating the need for trade-offs. For example, following the Dawes' rule (Dawes, 1979), individuals *simply* count the number of pros of each option (i.e., number of times the option performs better than its competitors), and then choose the one with the highest count. This rule simplifies the decision by effectively giving the same weight to each attribute (i.e., a particular case of equal AW).

Alternatively, individuals may restructure the multi-attribute information to decrease the number of trade-offs required to make a decision (i.e., moving from K to K' where $K' < K$). For example, non-compensatory decision rules in multi-attribute choices, such as ignoring attributes (Heidenreich et al., 2018; Hensher, 2006; Hole et al., 2013), or eliminating/selecting choice alternatives based on some decision criteria (e.g., elimination-by-aspects) (Erdem et al., 2014; Gilbride & Allenby, 2004; Tversky, 1972). Such rules imply discontinuities in the demand function, either precluding computation of MRS or leading to severe biases in the MRS when not accounted for (Campbell et al., 2011; Erdem et al., 2014; Heidenreich et al., 2018).

Our study contributes to the literature by testing an alternative rule referred to as attributes aggregation (AA). We propose a two-step decision-making process: individuals first aggregate attributes into different dimensions, referred to as "meta-attributes"; they then make trade-offs and choices based on these meta-attributes, referred to as "attributes aggregation" (AA). This proposed model can be seen as an extension of the Hierarchical Information Integration (HII) approach (Molin & Timmermans, 2009; Oppewal et al., 1994). HII has been applied to simplify the task when the choice experiment (CE) has a large number of attributes[1]; individuals divide attributes into subsets, then evaluate each subset separately and aggregate their evaluations to choose between competing options (Louviere & Timmermans, 1990). In our new model, individuals aggregate attributes and then trade across them.

A few studies in the health outcome literature have used techniques of restructuring information in their EQ-5D valuation exercises. For instance, the misery index (Ramos-Goñi et al., 2013) obtained by aggregating the five digits of an EQ-5D health state (e.g., health state 13255 is $1 + 3 + 2 + 5 + 5 = 16$), and counting the number of steps away from the perfect outcome (25), was used as a proxy for health states' severity levels. Shaw's US EQ-5D valuation (Shaw et al., 2005), using count levels, developed an ordinal variable, D1 (taking on values ranging from 0 to 4), that represented the number of movements away from perfect health (i.e., the additional number of dimensions at level 2 or level 3) beyond the first. In contrast to the AA model, both the misery index and Shaw's valuation specify a utility function with a constraint of equally-weighted attributes. Whilst AA is argued to be a simplifying decision rule, both the misery index and Shaw's D1 variable might be seen as making the decision rule more complex as individuals have to count how far they are from the "best" scenario, with "best" varying across individuals.

In the CE literature, AA as a decision rule has received very little attention. In the only study looking at this decision-making strategy when making multi-attribute choices, Layton and Hensher (2010) found that up to 88% of individuals aggregated attributes when making commuting decisions. More specifically, "*time spent in free flow traffic*" and "*time spent in congested traffic*" were translated into a new "*total traffic time*" attribute, and "*toll*" and "*running cost*" translated into a "*total cost*" attribute. Our study takes this initial work forward by developing a more general AA model. More specifically, we extend AA to the case of qualitative attributes which are frequently used in multi-attribute choices to describe nonmarket goods. We develop a model that allows categorical/qualitative attributes pertaining to the same construct to be jointly processed. We expect AA to be relevant in the health context, where attributes often share common features (e.g., process and outcome attributes, risk attributes, time attributes, qualitative attributes indicating health status severity and quality of care, and financial and nonfinancial incentives) (Clark et al., 2014; Soekhai et al., 2019). We explore different – and arguably more realistic – aggregation heuristics and examine the association

between survey/personal characteristics and AA to determine whether some individuals are more likely to use AA when making multi-attribute choices.

The rest of this paper is organized as follows. Section 2 describes the multi-attribute choice data. Section 3 describes the choice modelling approach, providing a benchmark model for comparison, and then describes the AA modelling, explaining its behavioral process. Section 4 presents the model results. Accounting for AA improves model fit, with the probability of adopting AA greater for homogenous information (consistent with our a priori hypothesis). AA is more prevalent: amongst those who failed the dominance test; when respondents make relatively faster choices; and when the choice tasks are placed at later positions. Section 5 presents the implications of AA for the monetary valuation of service improvements. Accommodating AA behavior leads to a reduction in the WTP estimates. Section 6 discusses the results and identifies avenues for future research. Section 7 draws conclusions.

2 | EXPERIMENTAL DESIGN AND SAMPLE

We used data from a CE concerned with preferences for personalization of chronic pain self-management programmes (Burton et al., 2017). We chose this dataset because we were extending AA to the case of qualitative attributes. In this CE, each choice option was described by four qualitative attributes: providing personalized information (INFORMA-TION); providing advice that matches personal situation (SITUATION); putting an emphasis on personal values in living well (LIVEWELL); and communication style (COMMUNICATION). In addition, a quantitative monetary attribute was included (COST). The attributes and their levels are shown in Table 1.

Twelve experimental choice tasks, each with three choice options, were generated using an S-efficient design (Bliemer & Rose, 2005; Rose et al., 2008) optimized for a multinomial logit (MNL) model. In each task, participants were asked to choose their preferred choice option (Figure 1). In addition to the 12 experimental tasks, each participant was asked two nonexperimental tasks which were manually added: a warm-up task (Task #1) to familiarize participants with the format of the choice questions and a dominance test (Task #14) (i.e., where one alternative in the choice set dominates all other alternatives). The questionnaire also included socio-demographic questions.

All respondents were provided with the 12 choice tasks. Extensive qualitative work indicated that individuals could manage 12 choice tasks. The order of the choice tasks was not randomized across respondents. The order of the choice options within a choice task was not randomized but was determined manually to ensure that option A was not always the most desirable option. The order of the attributes within the choice options was fixed (as in Figure 1).

The sample consisted of 517 members of a UK-based online panel managed by a research company, ResearchNow! (now called Dynata). These panel members were recruited via email and represented a diverse range

TABLE 1 Attributes and levels used to describe personalization of chronic pain self-management services

Attributes	Description	Levels
Information (INFORMATION)	Information about pain, the conditions that cause it, and the different ways there are of managing it	Provides everyone with the same information Provides information that is relevant to you[a]
Situation (SITUATION)	Things like where you live, who you live with, what resources you have, what you usually do for yourself and others, and how pain currently affects that	Takes little account of your current situation Makes suggestions that fit your current situation[a]
Living well (LIVING WELL)	Things that really matter to you, especially the kinds of things that you would like to achieve or to spend more time doing, and the kind of person that you want to be	Seems to think that everyone wants to get the same from life Works with you on what you want to get from life[a]
Communication (COMMUNICATION)	The way that the support service might communicate with you	Communicates with you in a neutral professional way Communicates with you in a friendly and personal way[a]
Cost (COST)		£5, £10, £15, £20

[a]Level corresponding to a higher level of personalization.

FIGURE 1 Illustration of a choice task used in the experiment [Colour figure can be viewed at wileyonlinelibrary.com]

of people in terms of socio-economic indicators and medical histories. Invitations were targeted to 16+ years old panel members diagnosed with chronic pain. The study was approved by the North of Scotland Research Ethics Service (Reference 14/NS/0075).

3 | ATTRIBUTES AGGREGATION IN MULTI-ATTRIBUTE CHOICES

3.1 | Random utility maximization

Discrete choices are typically modelled within the RUM framework, assuming random utility (Thurstone, 1927), multi-attributes utility (Lancaster, 1966), and utility maximization (Samuelson, 1938). The utility (U) of each choice option is decomposed into a systematic (V) and random (ε) component (Equation (1)). Following the LTD, V is defined by the product attributes (X) and their respective marginal utilities (β) (Equation (2)). Assuming ε is identically and independently distributed as type 1 extreme value (EV1) (Equation (3)), the choice probabilities (P) can be represented by the multinomial logit (MNL) model (Equation (5)) (McFadden, 1974; Train, 2009).

$$U_{ntj} = V_{ntj} + \varepsilon_{ntj} \tag{1}$$

$$V_{ntj} = f\left(X_{ntj}; \beta_{nk}\right) \tag{2}$$

$$\varepsilon_{ntj} \sim \text{iid EV1} \tag{3}$$

$$y_{nt} = i, \ \text{if} \ U_{nti} \geq U_{ntj}, \ \forall i \neq j \tag{4}$$

$$P(y_{nt} = i) = \frac{\exp(V_{nti})}{\sum_j \exp\left(V_{ntj}\right)} \tag{5}$$

U_{ntj} is the utility for participant n from alternative j in choice task t; β_k is the parameter to be estimated for the kth attribute; X_{ntjk} is the measured attribute k. y_{nt} is the choice made by n in t.

3.2 | Attribute weighting

In practice, the specification of the utility function typically takes the form of linear in parameters and additive in attributes, and then utility becomes a weighted average of the product attributes and preference parameters (Equation (6)). The standard formulation of the RUM implies that each product attribute is considered, such that it becomes possible to estimate individuals' sensitivities to marginal changes in the attributes (i.e., marginal utility).

$$V_{njt} = \sum_k \left(\beta_k X_{ntjk} \right) \tag{6}$$

As a reference model, we estimated an AW-MNL model assuming attributes weighting of the multi-attribute information. Assuming all five features are considered by the individuals when making a choice of chronic pain self-management programme, the AW-MNL model requires them to make ten trade-offs among the features (e.g., IN-FORMATION vs. SITUATION, INFORMATION vs. LIVEWELL, etc.,).

3.3 | Attribute aggregation

The AA can be represented by editing the indirect utility function (V). First, the individuals need to identify attributes which could be combined together. The four qualitative attributes describe the "*level of service personalization*" and then are conceptually related. Individuals can decide to aggregate these four attributes on the basis of their conceptual proximity (Equation (7)).

$$X_{ntjk}^{AA} = \left\{ INFORMATION_{ntj}, SITUATION_{ntj}, LIVEWELL_{ntj}, COMMUNICATION_{ntj} \right\} \tag{7}$$

Respondents will aggregate eligible attributes depending on their information structure (i.e., whether they provide similar or conflicting information about the good). For example, if two attributes describe a high level of personalization and the remaining two attributes describe a low level of personalization, then AA is unlikely to be relevant as the attributes provide mixed information about the service (i.e., the quality of the service is not unambiguously high or low). However, if the four qualitative attributes describe the same level of personalization, then the respondents are more likely to combine them into one single piece of information referred to as a meta-attribute (METAATTRIBUTE). Whilst it is possible to define AA as a very specific type of multi-way interaction effects between the attributes, it is unlikely to be implemented in practice as this would require including too many choice tasks in the CE to be practically feasible. We assumed that respondents evaluate the multi-attribute information by calculating a ratio of features[2] (φ) (Equation (8)).

$$\varphi_{ntj} = \frac{Min\left(COUNT_{ntj}^{High}, COUNT_{ntj}^{Neutral} \right)}{Max\left(COUNT_{ntj}^{High}, COUNT_{ntj}^{Neutral} \right)} \tag{8}$$

Equation (8) is a proxy for information homogeneity, calculating the number of attributes within a choice option with similar levels: a value closer to 1 indicates information heterogeneity, and nearer to 0 information homogeneity (i.e., attributes tend to share same levels)[3]. This ratio is based on the number of high personalization ($COUNT^{High}$) and neutral personalization ($COUNT^{Neutral}$) values.

Given the binary nature of the four qualitative attributes, the METAATTRIBUTE information is obtained by binary classification of the attributes following a majority rule (Equation (9)).

$$METAATTRIBUTE_{ntj} = \begin{cases} 1, & \text{if } \sum_k X^{AA}_{ntjk} > 2 \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

This aggregation rule gives the same importance to the four qualitative attributes in the aggregation process. This is consistent with the Dawes' rule following which individuals make choices by counting the number of positive/good features and selecting the option with the highest count (Dawes, 1979).

Respondents will then have to decide whether the information is sufficiently homogeneous to justify AA by comparing (φ) to a subjective threshold (α). If $\varphi \geq \alpha$, individuals retain the initial information structure (i.e., no attributes aggregation), and if $\varphi < \alpha$ individuals proceed to AA (Equation (10)).

$$U_{ntj} = \begin{cases} V^{AW}_{ntj} + \varepsilon_{ntj}, & \text{if } \varphi_{nt} \geq \alpha \\ V^{AA}_{ntj} + \varepsilon_{ntj}, & \text{if } \varphi_{nt} < \alpha \end{cases} \tag{10}$$

This initial formulation can be enriched by allowing the threshold value to differ across participants. For instance, some individuals may be willing to apply AA when at least half of the information (two attributes out of four) is similar, whilst some other participants would require all the information (four attributes out of four) to be similar. Similar to Layton and Hensher (2010), we allowed this threshold to be exponentially distributed in the sample, and then the probability of attributes aggregation (P^{AA}) becomes:

$$P^{AA}_{nt} = \exp(-\lambda \varphi_{nt}) \tag{11}$$

The λ parameter controls the degree of AA (i.e., λ close to zero implies strict AA).

Given the decision to apply AA depends on the structure of the multi-attribute information, it is likely to be made at the task-level (i.e., a given individual can decide to apply AA only in some of the choice tasks). This also implies that individuals would restructure the multi-attribute information for none or all of the options included in the task[4]. Individuals were assumed to evaluate the homogeneity of the multi-attribute information for each option (A, B, and C) separately, leading thus to three measures (φ_A, φ_B, φ_C), and then would use them to make a joint decision of AA.

$$P_{nt}(AA) = P(\varphi_{ntA} < \alpha_n) \times P(\varphi_{ntB} < \alpha_n) \times P(\varphi_{ntC} < \alpha_n) = \exp\left(-\lambda \sum_j \varphi_{ntj}\right) \tag{12}$$

Following this conceptualization of AA, we estimated the following AA-MNL model:

$$\begin{aligned} U_{ntj} = & \left(\beta_1 INFORMATION_{ntj} + \beta_2 SITUATION_{ntj} + \beta_3 LIVEWELL_{ntj} + \beta_4 COMMUNICATION_{ntj}\right)\left(1 - P^{AA}_{nt}\right) \\ & + \beta METAATTRIBUTE_{ntj}\left(P^{AA}_{nt}\right) + \delta_1 ASC1_{ntj} + \delta_2 ASC2_{ntj} + \gamma COST_{ntj} + \varepsilon_{ntj} \end{aligned} \tag{13}$$

We then allowed the threshold value (λ) to depend on the characteristics of both the respondents and choice task.

$$\lambda_n = \mu_1 EDUCATION_n + \mu_2 DOMINANCE_n + \mu_3 RESPONSETIME_{nt} + \mu_4 LOCATION_n \tag{14}$$

where EDUCATION is education (capturing the effect of university/college education compared to less than college or other education), DOMINANCE refers to whether the respondent passes the dominance test[5] (see Section 2), RESPONSETIME is the response time (RT) measured in seconds at the task level, and LOCATION is the location of the choice tasks in the questionnaire (Table 2).

It is argued that education develops one's ability to gather and interpret information, increasing one's ability to solve difficult problems (Ross & Chia-Ling Wu, 1995). We hypothesize that respondents with a higher level of education would be less likely to adopt AA as an information processing rule.

TABLE 2 Survey and participant-level characteristics (N = 517)

		%
EDUCATION	Less than University/other	57.4
	University	42.6
Choice task LOCATION	Block 1 (tasks 1–4)	33.3
	Block 2 (tasks 5–8)	33.3
	Block 3 (tasks 9–12)	33.3
RESPONSETIME	Faster RT	20.59
	Normal RT	31.27
	Slower RT	48.14
DOMINANCE test	Pass	86.46
	Fail	13.54

Note: For RESPONSETIME our Quartiles are for the RT distribution for each choice task, not the distribution of the population by time. For each choice task, we first computed the first and third quartiles of the RT distribution (i.e., the 0.25 quantile and 0.75 quantile RT). A RT is classified as a fast response when RT is less than or equal to the first quartile and a slow response when RT is greater than the third quartile of the distribution. Based on this definition, 20% of responses were classified as "fast" and 48% as "slow".

Abbreviation: RT, response time.

In line with San Miguel et al. (2005), we expect respondents who failed the dominance test to be more likely to find the choice tasks difficult, less likely to give attention to each attribute of the good and hence more likely to adopt a choice simplifying rule while completing choice tasks. We thus hypothesize that AA is more likely to occur among individuals who failed the dominance test.

While it is possible that respondents who answered relatively quickly processed all the information in the choice tasks and made a utility-maximizing choice (Börger, 2016), it is also likely that they utilized AA as a decision-making heuristic. Holmes et al. (1998) find that respondents who took little time to answer multi-attribute choices did not respond in ways that conform to underlying economic theory. We test the hypothesis that respondents that rush through the experiment (faster RT) may not sufficiently consider all information provided and hence are more likely to aggregate attributes. For each choice task, we first computed the first and third quartiles of the RT distribution. An RT is classified as a fast response when its duration is less than or equal to the first quartile and a slow response when its duration is greater than the third quartile.

The literature suggests that individuals may exhibit two forms of heterogeneity (learning and fatigue) within the sequence of their choices (Campbell et al., 2015) and that such forms of heterogeneity within task sequence is likely to affect the adoption of AA. Braga and Starmer (2005) identified two forms of learning within a valuation context: institutional learning whereby individuals learn the rules of the market (real or hypothetical) and value learning whereby individuals gain knowledge of their preferences for the good under investigation. Moreover, over a sequence of choices, asking respondents to make a large number of complex choices make them fatigued or bored and increasingly confused (Alberini, 2012). Hess et al. (2012) noted that later located choice tasks will reflect a higher dimension of variability. Swait and Adamowicz (2001) indicated an inverted bell-shaped effect of repeated task position on consistency, reflecting *learning effects* for an early position of the repeated choice task, and *fatigue effects* for later positions. In this regard, we hypothesize that respondents may be less likely to adopt AA as a decision rule in the first few choice tasks (reflecting *learning effects* for an early position of the repeated choice task), but as they progress through the choice tasks, a fatigue or boredom effect can make them more likely to adopt AA (*fatigue effects* for later positions). Given we did not randomly order the tasks, including LOCATION also allowed us to control for task sequence to capture the effect of task location/order on AA.

4 | RESULTS

The results for the reference AW-MNL model is presented in Table 3, column 2. We also estimated an error component logit (ECL) model, dropping the independence of irrelevant alternatives (IIA) assumption. However, this model failed to

outperform the MNL model, but estimation time increased. Using a log-likelihood ratio test (LR test: Deviance = 2.2; dof = 3; p-value = 0.532), we found no evidence of differences in respondents' choices between the two models. We thus focus on the MNL model. All coefficients were significant in the expected directions (i.e., a positive effect for improvement in the personalization dimensions and a negative effect for a COST increase).

Results of the AA-MNL model are presented in Table 3, column 3. We obtained a similar pattern of preferences for all five attributes. The coefficient of the METAATTRIBUTE is positive and significant, implying participants prefer a higher improvement in the service personalization. The λ parameter was statistically significant, indicating that our model identified 21.5% ($n = 111$)[6] of the participants as "attributes aggregators".

Figure 2 displays the variation in the probability of AA when information heterogeneity changes, *ceteris paribus*. The negatively sloped part of the plot indicates that the probability of AA declines as information heterogeneity increases. For a very low level of information heterogeneity (for instance, $\varphi = 0.1$), there is a 0.9 probability of AA, suggesting that more homogenous information is likely to be aggregated.

Accounting for task and individuals' characteristics further improved modelling performance, as indicated by the lower likelihood (LL) value (LL$_{AA \text{ with heterogeneity}} = -5731.6$ vs. LL$_{AA} = -5856$ vs. LL$_{AW} = -5879.7$; Table 3, column 4). Using the LR test between the restricted (AW) and unrestricted (AA) models, the AA model provides a significantly better fit compared to the AW model (dev = 40, dof = 8, $p < 0.001$). We also compared the models based on the adjusted McFadden's pseudo-R^2 (Mokhtarian, 2016) to give an idea of modelling performance: accounting for AA results in a modest improvement in model fit[7] (adjusted McFadden's pseudo R^2: AW-MNL = 0.129; AA-MNL = 0.132; AA-MNL $_{\text{with heterogeneity}}$ = 0.150).

TABLE 3 Results of MNL models allowing for attributes aggregation

	AW-MNL MLE (SE)	AA-MNL MLE (SE)	AA-MNL with covariates MLE (SE)
INFORMATION (β_1)	0.676 (0.028)***	0.801 (0.059)***	1.000 (0.083)***
SITUATION (β_2)	0.959 (0.039)***	1.183 (0.080)***	1.604 (0.132)***
LIVEWELL (β_3)	0.786 (0.034)***	0.938 (0.063)***	1.341 (0.104)***
COMMUNICATION (β_4)	0.249 (0.028)***	0.242 (0.052)***	0.455 (0.055)***
COST (γ)	−0.053 (0.003)***	−0.048 (0.003)***	−0.052 (0.003)***
ASC1 (δ_1)	0.165 (0.035)***	0.121 (0.036)***	0.049 (0.038)
ASC2 (δ_2)	0.005 (0.034)	−0.005 (0.035)	−0.023 (0.035)
METAATTRIBUTE	-	0.975 (0.216)***	0.062 (0.119)
Extent of aggregation (λ)	-	1.040 (0.169)***	1.148 (0.301)***
EDUCATION (University)	-	-	0.016 (0.028)
DOMINANCE test (fail)	-	-	−0.084 (0.029)***
Task LOCATION 2 (tasks 5–8)	-	-	0.187 (0.059)***
Task LOCATION 3 (tasks 9–12)	-	-	−0.203 (0.044)***
RESPONSETIME (faster)	-	-	−0.953 (0.284)***
RESPONSETIME (slower)	-	-	0.987 (0.345)***
Number of observations	6204	6204	6204
Sample size (N)	517	517	517
Number of parameters	7	9	15
Log-likelihood (LL)	−5879.7	−5856	−5731.6
Adjusted McFadden's pseudo R^2	0.129	0.132	0.150

Abbreviations: AA, attribute aggregation; ASC, alternative specific constant; AW, attributes weighting; MLE, maximum likelihood estimator; MNL, multinomial logit; SE, standard error.
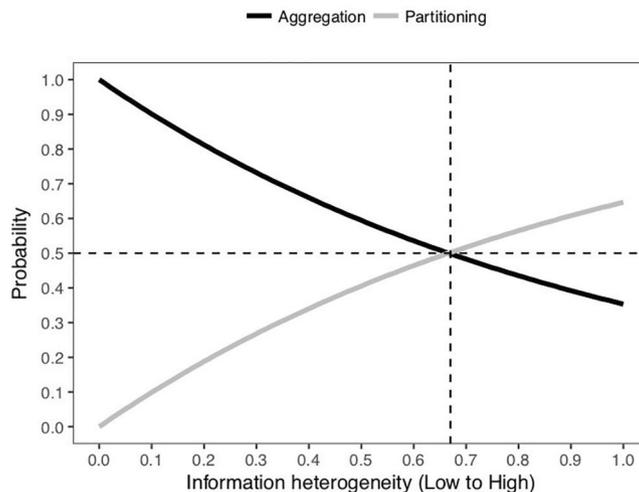
*** indicates significance at 1% level.

FIGURE 2 Cumulative density functions of attributes weighting versus attributes aggregation for a range of information heterogeneity

Respondents who failed the DOMINANCE test had a lower threshold and hence were more likely to adopt AA. Failing the DOMINANCE test, therefore, may imply difficulty of the choice tasks and/or less attention to carefully consider each piece of information and hence more likely to use AA as a choice simplifying mechanism. The LOCATION of the choice tasks also affects the aggregation threshold; compared to the first four choice tasks, respondents are less likely to adopt AA for the middle-located tasks (tasks 5–8) but more likely to aggregate attributes when the choice tasks are located in the later positions (tasks 9–12). A shorter RESPONSETIME affected the threshold negatively, meaning respondents who spent a shorter time completing the tasks are more likely to aggregate attributes (compared to those who spent a relatively longer time). We find evidence of a strong relationship between response time and probability of adopting AA. For instance, for an information heterogeneity equal to one ($\varphi = 1$), the average probability of AA is 32%, but when choices are made quickly, this probability goes up to 82% and while it decreases to 12% when respondents took longer time to complete choices, suggesting a very strong relationship between RESPONSETIME and AA. EDUCATION had no effect on the probability of adopting AA.

5 | IMPLICATIONS OF AA FOR MEASUREMENT OF WILLINGNESS TO PAY VALUES

When analyzing multi-attribute choice data, willingness to pay (WTP) can be estimated as the MRS between a given attribute (k) and the COST attribute (Equation (15)).

$$\text{WTP}_k = -\frac{\partial V_{ntj} / \partial X_{ntj(k)}}{\partial V_{ntj} / \partial X_{ntj(\text{COST})}} \tag{15}$$

As COST was not aggregated, the marginal utility of COST is the same in both AW and AA models: $\frac{\partial V_{ntj}}{\partial X_{ntj(\text{COST})}} = \gamma$

However, the marginal utility of the other (X_k) attribute differs across the model specifications. In the AW model:

$$\partial V_{ntj} / \partial X_{ntj(k)} = \beta_k$$

Whereas in the AA model:

$$\frac{\partial V_{ntj}}{\partial X_{ntj(k)}} = \beta_k (1 - e^{-\lambda \varphi_{nt}}) \tag{16}$$

$$\frac{\partial V_{ntj}}{\partial \text{META}} = \beta e^{-\lambda \varphi_{nt}} \tag{17}$$

TABLE 4 Willingness to pay (WTP, £) estimates for high levels of personalization

Attributes	AW-MNL	AA-MNL				
		$\varphi = 0.1$	$\varphi = 0.3$	$\varphi = 0.5$	$\varphi = 0.7$	$\varphi = 0.9$
INFORMATION	12.755	1.648	4.473	6.766	8.630	10.143
SITUATION	18.094	2.434	6.606	9.993	12.745	14.980
LIVEWELL	14.830	1.930	5.238	7.924	10.106	11.878
COMMUNICATION	4.698	0.498	1.351	2.044	2.607	3.064
METAATTRIBUTE	-	18.306	14.868	12.076	9.808	7.966

Note: φ: The ratio of features measure indicating the extent of information homogeneity/heterogeneity. The values are arbitrarily chosen in the range of zero and one. $\varphi = 0.1$: Information is less heterogenous. $\varphi = 0.9$: Information is more heterogenous.

Abbreviations: AA, attribute aggregation; AW, attributes weighting; MNL, multinomial logit.
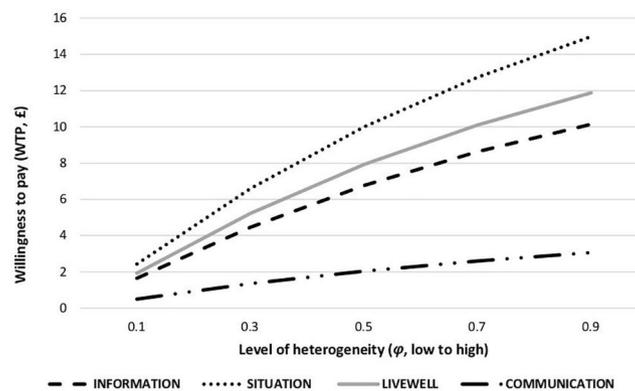

FIGURE 3 Willingness to pay for high levels of personalization at various levels of information heterogeneity

To use the WTP formula for the AA model, we assign a specific value for φ. We considered five arbitrary[8] values, between 0 and 1. Results are presented in Table 4 and Figure 3. WTP for personalization attributes derived from the AW model is higher compared to the AA model, though the difference reduces for higher levels of information heterogeneity (e.g., $\varphi_{nt} = 0.9$). When attribute information become less heterogeneous (e.g., $\varphi_{nt} = 0.1$), the differences in WTP between AW-MNL and AA-MNL gets bigger for all personalization attributes (Table 4) suggesting a positive relationship between information heterogeneity and WTP for higher levels of each personalization attribute.

6 | DISCUSSION

Economic analysis of multi-attribute choices adopts an extreme version of the Lancasterian theory of demand (LTD), assuming attributes exert a direct influence on individuals' choices. The behavioral realism of this assumption is debatable. Our attribute aggregation (AA) model allows individuals to *translate* the multi-attribute information into meta-attributes. We test the empirical validity of our AA model using data from a multi-attribute CE concerned with patients' preferences for delivery of chronic pain management services. Five attributes described the multi-attribute choice options; four qualitative attributes were conceptually related (degree of personalization), and the 5th attribute was cost. We translated the four qualitative attributes into a meta "personalization" attribute. Approximately twenty percent of respondents were attribute aggregators. Respondents who make relatively fast and/or illogical choices were more likely to be attribute aggregators, suggesting AA could be related to a lower level of engagement in the decision-making process.

Allowing for AA had a significant impact on WTP values. Our results may help explain the observed "part-whole bias" in the monetary valuation of public goods. Bateman et al. (1997) showed that if components are evaluated separately, the sum of those valuations exceeds the value placed overall. We found a similar result: the sum of WTP values for the four qualitative attributes was £48 compared to £18 when modelled as a meta-attribute.

Our result may also be linked to "support theory", a psychological model of a degree of belief, which assumes that the judged probability of an event increases when its description is unpacked into disjoint components. Rottenstreich and Tversky (1997) showed that when individuals are presented with an explicit disjunction (for instance, the probability that a particular student specializes in health economics, environmental economics, or agricultural economics), they may repack the various disciplines and evaluate the whole component 'economics' rather than the separate specializations. The authors note the presence of more explicit additivity for similar components than dissimilar components because similar parts are more easily repacked.

Our study could be extended in several ways. Estimation of the aggregation threshold (α) took place at the respondent level. One could investigate changes in the threshold across choice tasks (i.e., $\alpha_n \rightarrow \alpha_{nt}$), allowing for dynamic changes in decision-making, that is, participants may be less likely to adopt AA as a decision rule in the first few choice tasks, but as they go through the sequence of tasks, a fatigue or boredom effect can make them more likely to adopt AA and then to lower the threshold value. Second, the aggregation rule adopted gave the same importance to the four qualitative attributes, consistent with the Dawes' rule (Dawes, 1979). Future studies could make use of self-reported information about attributes importance to refine the weighting scheme. We have assumed AA is easier than the AW model, and our quantitative analysis supported these results[9]. We used the ratio of features to determine the percentage of attributes with similar levels (as a proxy for information homogeneity). This rule requires attributes to have the same format. Whilst Layton and Hensher's approach enables the aggregation of numerical attributes (Layton & Hensher, 2010), our approach allows the aggregation of qualitative attributes. We leave aggregation of both numerical and qualitative attributes to future research.

The CE used in this paper was not designed to test for AA. The reason we chose the preference for personalized care study is that we hypothesized that given the nature of the attributes, AA was likely. Further, we were extending the AA model to the case of qualitative variables. We recognize that the format is not typical of many CEs. Indeed, given that the format is argued to be easier for respondents to answer (Lancsar et al., 2013), our finding of AA might be argued to be stronger. However, we suggest future research explores our AA model using other CE formats. We reran our AA model on two additional data set. Our first data set was a CE which adopted a more standard forced choice to elicit patient preferences for kidney transplantations (Genie et al., 2020). The results, available in the Online Supplementary Material, supported the AA model. Secondly, the preference for personal care study included a follow-up question asking respondents if they would buy their preferred option. We reran the analysis on this data, including the opt-out option. Results again supported the AA model and are available from the authors. Whilst our analysis supports the AA model, we do not assume that AA always exists, but that it is a decision strategy that should be explored at the analytical stage of the CE. Future research could explore using other research methods, such as think-aloud (Ryan et al., 2009) and debriefing questions (Layton & Hensher, 2010; Pearce et al., 2020) concerning AA (i.e., whether respondents aggregate attributes or not).

This study is not exempt from limitations. It might be argued that AA is the result of a poorly developed CE, with the specified utility function including attributes that were either irrelevant or not well defined. In our specific case, this is unlikely: we developed the CE survey using good practice guidelines with extensive qualitative developmental work. There is a risk of a confounding effect between AA and preference heterogeneity. For example, if a respondent's preferences for the four qualitative attributes are similar, our model would erroneously consider this pattern of preferences as AA. However, we note there is great heterogeneity in individuals' choices, suggesting this confounding effect will be minimal. Future research could explore combining our AA approach with more flexible choice models such as mixed or latent class logit (Hensher & Greene, 2003; Shen, 2009). Our AA model relies on a number of sensible but arbitrary assumptions. Further research should explore the sensitivity of our results to the choice of the cut-off point or the form of aggregation (simple arithmetic vs. binary classification of information for different cut-off points). We recognize that CEs are increasingly using block designs to reduce the task burden, with choices displayed randomly during the experiment. Given our data set did neither of these it might be argued that observed AA was a result of fatigue and the difficulty of choice.[10] Given the easier format of the CE approach adopted, and the detailed developmental work, this is unlikely to be the case. Further, whilst we did not randomize the order in which choices were presented, we controlled for task sequence to capture the effect of task location on attributes aggregation.

7 | CONCLUSION

Our results underline the importance of accounting for information processing rules when modelling multi-attribute choices. More specifically, we provide evidence for further inquiry into the use of AA when responding to CEs.

Accounting for AA has implications for welfare estimates. Future research should replicate this approach on CE data sets using different attributes and choice formats and apply other methods (e.g., think-aloud methods and debriefing questions). Such research will increase the validity of welfare estimates generated.

DATA AVAILABILITY STATEMENT
Research data are not shared due to privacy/ethical restrictions.

ORCID
Mesfin G. Genie https://orcid.org/0000-0002-1744-4666
Nicolas Krucien https://orcid.org/0000-0001-9179-6558
Mandy Ryan https://orcid.org/0000-0001-6643-9129

ENDNOTES
[1] For example, in a preference study concerned with breast cancer surgery guidelines, HII was integrated with CE (van Helvoort-Postulart et al., 2009) to create three constructs from 15 attributes: (i) *Organization* ("day surgery unit," "breast-care nursing staff," "compensation," "discharge criteria," and "collaboration agreements with home care organizations"); (ii) *Cooperation partners* ("patients/patient organizations," "colleagues," "management," "ward nurses," and "expertise of home care nurses"); and (iii) *Patient-centredness* ("written information after diagnosis," "preoperative counselling," "written information at discharge," "possibility to choose between day-care and hospital admission," and "patient satisfaction").

[2] We also used standard deviation (SD) of the attributes' levels as an alternative measure of information homogeneity, but the corresponding AA-MNL model was associated with a lower level of statistical performance. The results are available up on request.

[3] For example, based on the choice task in Figure 1, each qualitative attribute in Service A is described as [High, Neutral, Neutral, Neutral]; Service B as [High, High, High, High]; and Service C as [Neutral, Neutral, High, High]. Using Equation (8), the count of "High" for Service A is one and the count of "Neutral" is three; hence Min (1, 3) = 1, that is, the minimum of the combination is 1. The maximum of the combination, Max (1, 3) = 3; the ratio of the two (1/3) = 0.3, indicating more homogenous information. For Service B, the count of "High" is four and the count of "Neutral" is zero; hence Min (4, 0) = 0, that is, the minimum of the combination is 0, and the maximum of the combination Max (4, 0) = 4. Hence, the ratio of features becomes (0/4) = zero, suggesting that the four attributes provide homogenous information. The opposite is found for Service C, with two attributes taking "High" values and two attributes "Neutral" values. With Min (2, 2) = Max (2, 2), the ratio = 1, showing attributes provide heterogeneous information. Perfect information homogeneity occurs when a service is characterized by either 4 High & 0 Neutral or 4 Neutral and 0 High levels (i.e., ratio = 0). Perfect heterogeneity occurs when a service is described by 2 Neutral & 2 High (i.e., ratio = 1).

[4] This approach prevents respondents applying AA for one option and AW for another, as this would imply some forms of asymmetric comparisons (e.g., METAATTRIBUTE vs. INFORMATION) which are behaviorally difficult to justify.

[5] Dominance tests are included to check whether participants hold monotonic preferences; that is, they choose options where one alternative is more attractive than the other on all features.

[6] To compute the number of attribute aggregators in the sample, we followed the following steps: first, we calculate the AA probability for each task and individual (using Equation (14)); second, we compute the average AA probability for each individual; and finally, an individual is classified as an "aggregator" when the average AA probability is greater or equal to 50%. Then, we count those individuals whose average AA probability is greater than or equal to 0.5 (50%).

[7] We also compared the standard AW and AA models in terms of predictive performance. The predictive performance of each approach was computed as the percentage of observed choices correctly predicted by the model. We then used mean score to compare the two approaches. Both the AW and AA approaches perform equally in terms of predictive performance (AW: 67.8% [95% CI: 65.5–70.2]; AA: 67.4% [95% CI: 64.2–70.5]).

[8] Any other values between zero and one can be possible. The values 0.1, 0.3, 0.5, 0.7 and 0.9 are chosen arbitrarily to check the changes in WTP as information heterogeneity changes. A value of 0.9 means that information is more heterogeneous compared to the other values.

9 A CE requires respondents to make two cognitively demanding operations: (i) process the multi-attribute information (i.e., look at the attributes and understand their meaning, and extract value from the information) and (ii) make comparisons across choice options. Previous studies have shown that making trade-offs is very difficult for most people (Luce et al., 1999; Retief et al., 2013). By decreasing the number of trade-offs, AA should, in principle, decrease the cognitive difficulty of the choice tasks. Suppose, for instance, respondents make a choice between 2 options (A vs. B), each described by four attributes (1–4). Under AW, people would need to make 16 pairwise comparisons (A1 vs. B1; A1 vs. B2; etc.). Under AA, and assuming that attributes 1 and 2 belong to dimension D1 and attributes 3 and 4 belong to dimension D2, people would need to make only four comparisons (AD1 vs. BD1; AD1 vs. BD2; etc.).

10 We thank an anonymous referee for making this point.

REFERENCES

Alberini, A. (2012). Repeated questioning in choice experiments: Aare we improving statistical efficiency or getting respondents confused? *Journal of Environmental Economics and Policy*, *1*(2), 216–233. https://doi.org/10.1080/21606544.2012.698516

Araña, J. E., & León, C. J. (2009). Understanding the use of non-compensatory decision rules in discrete choice experiments: The role of emotions. *Ecological Economics*, *68*(8–9), 2316–2326. https://doi.org/10.1016/j.ecolecon.2009.03.003

Bateman, I., Munro, A., Rhodes, B., Starmer, C., & Sugden, R. (1997). Does part-whole bias exist? An experimental investigation. *Economic Journal*, *107*(441), 322–332. https://doi.org/10.1111/j.0013-0133.1997.160.x

Ben-Akiva, M. E., Lerman, S. R., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand* (Vol. 9). MIT press.

Bliemer, M. C. J., & Rose, J. M. (2005). Efficiency and sample size requirements for stated. Transport, ITLS-WP-05 (April).

Börger, T. (2016). Are fast responses more random? Testing the effect of response time on scale in an online choice experiment. *Environmental and Resource Economics*, *65*(2), 389–413. https://doi.org/10.1007/s10640-015-9905-1

Borghans, L., Romans, M., & Sauermann, J. (2010). What makes a good conference? Analysing the preferences of labour economists. *Labour Economics*, *17*(5), 868–874. https://doi.org/10.1016/j.labeco.2010.04.009

Braga, J., & Starmer, C. (2005). Preference anomalies, preference elicitation and the discovered preference hypothesis. *Environmental and Resource Economics*, *32*(1), 55–89. https://doi.org/10.1007/s10640-005-6028-0

Breustedt, G., Müller-scheeßel, J., & Latacs-Lohman, U. (2008). Forecasting the adoption of GM oilseed rape: Evidence from a discrete choice experiment in Germany. *Journal of Agricultural Economics*, *59*(2), 237–256. https://doi.org/10.1111/j.1477-9552.2007.00147.x

Burton, C. D., Entwistle, V. A., Elliott, A. M., Krucien, N., Porteous, T., & Ryan, M. (2017). The value of different aspects of person-centred care: A series of discrete choice experiments in people with long-term conditions. *BMJ Open*, *7*(4). e015689. https://doi.org/10.1136/bmjopen-2016-015689

Campbell, D., Boeri, M., Doherty, E., & George Hutchinson, W. (2015). Learning, fatigue and preference formation in discrete choice experiments. *Journal of Economic Behavior & Organization*, *119*, 345–363. https://doi.org/10.1016/j.jebo.2015.08.018

Campbell, D., Hensher, D. A., & Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: A latent class specification. *Journal of Environmental Planning and Management*, *54*(8), 1061–1076. https://doi.org/10.1080/09640568.2010.549367

Clark, M. D., Determann, D., Petrou, S., Moro, D., & de Bekker-Grob, E. W. (2014). Discrete choice experiments in health economics: A review of the literature. *PharmacoEconomics*, *32*(9), 883–902. https://doi.org/10.1007/s40273-014-0170-x

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*(7), 571–582. https://doi.org/10.1037/0003-066X.34.7.571

de Bekker-Grob, E. W., Ryan, M., & Gerard, K. (2012). Discrete choice experiments in health economics: A review of the literature. *Health Economics*, *21*(2), 883–902. https://doi.org/10.1002/hec.1697

Erdem, S., Campbell, D., & Thompson, C. (2014). Elimination and selection by aspects in health choice experiments: Prioritising health service innovations. *Journal of Health Economics*, *38*,10–22. https://doi.org/10.1016/j.jhealeco.2014.06.012

Genie, M. G., Nicoló, A., & Pasini, G. (2020). The role of heterogeneity of patients' preferences in kidney transplantation. *Journal of Health Economics*, *72*. 102331. https://doi.org/10.1016/j.jhealeco.2020.102331

Gilbride, T. J., & Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, *23*(3), 391–406. https://doi.org/10.1287/mksc.1030.0032.

Hackbarth, A., & Madlener, R. (2013). Consumer preferences for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment*, *25*, 5–17. https://doi.org/10.1016/j.trd.2013.07.002

Hansen, F. (1976). Psychological theories of consumer choice. *Journal of Consumer Research*, *3*(3), 117–142. https://doi.org/10.1086/208660

Heidenreich, S., Watson, V., Ryan, M., & Phimister, E. (2018). Decision heuristic or preference? Attribute non-attendance in discrete choice problems. *Health Economics*, *27*(1), 157–171. https://doi.org/10.1002/hec.3524

Hensher, D. A. (1994). Stated preference analysis of travel choices: The state of practice. *Transportation*, *21*(2), 107–133. https://doi.org/10.1007/BF01098788

Hensher, D. A. (2006). How do respondents process stated choice experiments? Attribute consideration under varying information load. *Journal of Applied Econometrics*, *21*(6), 861–878. https://doi.org/10.1002/jae.877

Hensher, D. A., & Greene, W. H. (2003). The mixed logit model: The state of practice. *Transportation*, *30*(2), 133–176. https://doi.org/10.1023/A:1022558715350

Hess, S., Hensher, D. A., & Daly, A. (2012). Not bored yet – Revisiting respondent fatigue in stated choice experiments. *Transportation Research Part A: Policy and Practice*, *46*(3), 626–644. https://doi.org/10.1016/j.tra.2011.11.008

Hole, A. R., Kolstad, J. R., & Gyrd-Hansen, D. (2013). Inferred vs. stated attribute non-attendance in choice experiments: A study of doctors' prescription behaviour. *Journal of Economic Behavior & Organization*, *96*, 21–31. https://doi.org/10.1016/j.jebo.2013.09.009

Holmes, T., Alger, K., Zinkhan, C., & Mercer, E. (1998). The effect of response time on conjoint analysis estimates of rainforest protection values. *Journal of Forest Economics*, *4*(1), 7–28.

Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics*. *69*(8), 1595–1603. https://doi.org/10.1016/j.ecolecon.2010.04.011

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157. https://doi.org/10.1086/259131

Lancsar, E., & Louviere, J. (2008). Conducting discrete choice experiments to inform healthcare decision making: A user's guide. *PharmacoEconomics*, *26*(8), 661–677. https://doi.org/10.2165/00019053-200826080-00004

Lancsar, E., Louviere, J., Donaldson, C., Currie, G., & Burgess, L. (2013). Best worst discrete choice experiments in health: Methods and an application. *Social Science & Medicine*, *76*(1), 74–82. https://doi.org/10.1016/j.socscimed.2012.10.007

Layton, D., & Hensher, D. A. (2010). Aggregation of common-metric attributes in preference revelation and implications for willingness to pay. *Transportation Research Part D: Transport and Environment*, *15*(7), 394–404. https://doi.org/10.1016/j.trd.2010.04.007

Louviere, J. J., & Timmermans, H. J. P. (1990). Using hierarchical information integration to model consumer responses to possible planning actions: Recreation destination choice illustration. *Environment & Planning A*, *22*(3), 291–308. https://doi.org/10.1068/a220291

Louviere, J. J., Hensher, D. A., Swait, J. D., & Adamowicz, W. (2000). *Stated choice methods: Analysis and application*. Cambridge University Press.

Luce, M. F., Payne, J. W., & Bettman, J. R. (1999). Emotional trade-off difficulty and choice. *Journal of Marketing Research*, *36*(2), 143–159. https://doi.org/10.2307/3152089

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In *Frontier in econometrics* (pp. 105–142). New York: Academic Press.

Mokhtarian, P. L. (2016). Discrete choice models' ρ2: A reintroduction to an old friend. *Journal of Choice Modelling*, *21*, 60–65. https://doi.org/10.1016/j.jocm.2016.02.001

Molin, E. J. E., & Timmermans, H. J. P. (2009). Hierarchical information integration experiments and integrated choice experiments. *Transport Reviews*, *29*(5), 635–655. https://doi.org/10.1080/01441640902829470

Oppewal, H., Louviere, J. J., & Timmermans, H. J. (1994). Modelling hierarchical conjoint processes with integrated choice experiments. *Journal of Marketing Research*, *31*(1), 92–105.

Pearce, A. M., Mulhern, B. J., Watson, V., & Viney, R. C. (2020). How are debriefing questions used in health discrete choice experiments? An online survey. *Value in Health*, *23*(3), 289–293. https://doi.org/10.1016/j.jval.2019.10.001

Ramos-Goñi, J. M., Rivero-Arias, O., Errea, M., Stolk, E. A., Herdman, M., & Cabasés, J. M. (2013). Dealing with the health state "dead" when using discrete choice experiments to obtain values for EQ-5D-5L heath states. *The European Journal of Health Economics*, *14*(SUPPL. 1), S33–S42. https://doi.org/10.1007/s10198-013-0511-2

Retief, F., Morrison-Saunders, A., Geneletti, D., & Pope, J. (2013). Exploring the psychology of trade-off decision-making in environmental impact assessment. *Impact Assessment and Project Appraisal*, *31*(1), 13–23. https://doi.org/10.1080/14615517.2013.768007

Rose, J. M., Bliemer, M. C. J., Hensher, D. A., & Collins, A. T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological*, *42*(4), 395–406. https://doi.org/10.1016/j.trb.2007.09.002

Ross, C. E., & Wu, C.-L. (1995). The links between education and health. *American Sociological Review*, *60*(5), 719–745. https://doi.org/10.2307/2096319

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review*, *104*(2), 406–415. https://doi.org/10.1037/0033-295X.104.2.406

Ryan, M., Watson, V., & Entwistle, V. (2009). Rationalising the "irrational": A think aloud study of discrete choice experiment responses. *Health Economics*, *18*(3), 321–336. https://doi.org/10.1002/hec.1369

Ryan, M., Gerard, K., & Amaya-Amaya, M. (Eds.). (2007). *Using discrete choice experiments to value health and health care* (Vol. *11*). Springer Science & Business Media.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, *5*(17), 61–71. https://doi.org/10.2307/2548836

San Miguel, F., Ryan, M., & Amaya-Amaya, M. (2005). "Irrational" stated preferences: A quantitative and qualitative investigation. *Health Economics*, *14*(3), 307–322. https://doi.org/10.1002/hec.912

Scott, A. (2002). Identifying and analysing dominant preferences in discrete choice experiments: An application in health care. *Journal of Economic Psychology*, *23*(3), 383–398. https://doi.org/10.1016/S0167-4870(02)00082-X

Shaw, J. W., Johnson, J. A., & Coons, S. J. (2005). US valuation of the EQ-5D health states: Development and testing of the D1 valuation model. *Medical Care*, *43*(3), 203–220. https://doi.org/10.1097/00005650-200503000-00003

Shen, J. (2009). Latent class model or mixed logit model? A comparison by transport mode choice data. *Applied Economics*, *41*(22), 2915–2924. https://doi.org/10.1080/00036840801964633

Simon, H. (1957). *A behavioural model of rational choice, in models of man, social and rational: Mathematical essays on rational human behaviour in a social setting*. Wyley.

Soekhai, V., de Bekker-Grob, E. W., Ellis, A. R., & Vass, C. M. (2019). Discrete choice experiments in health economics: Past, present and future. *PharmacoEconomics*, *37*(2), 201–226. https://doi.org/10.1007/s40273-018-0734-2

Swait, J., & Adamowicz, W. (2001). Choice environment, market complexity, and consumer behaviour: A theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organizational Behavior and Human Decision Processes*, *86*(2), 141–167.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *34*(4), 273–286. https://doi.org/10.1037/h0070288

Train, K. E. (2009). Discrete choice methods with simulation. In *Discrete choice methods with simulation* (2nd ed., Vol. 9780521766555). Cambridge University Press. https://doi.org/10.1017/CBO9780511805271

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299. https://doi.org/10.1037/h0032955

van Helvoort-Postulart, D., Dellaert, B. G. C., van der Weijden, T., von Meyenfeldt, M. F., & Dirksen, C. D. (2009). Discrete choice experiments for complex health-care decisions: Does hierarchical information integration offer a solution? *Health Economics*, *18*(8), 903–920. https://doi.org/10.1002/hec.1411

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Genie MG, Krucien N, Ryan M. Weighting or aggregating? Investigating information processing in multi-attribute choices. *Health Economics*. 2021;30:1291–1305. https://doi.org/10.1002/hec.4245