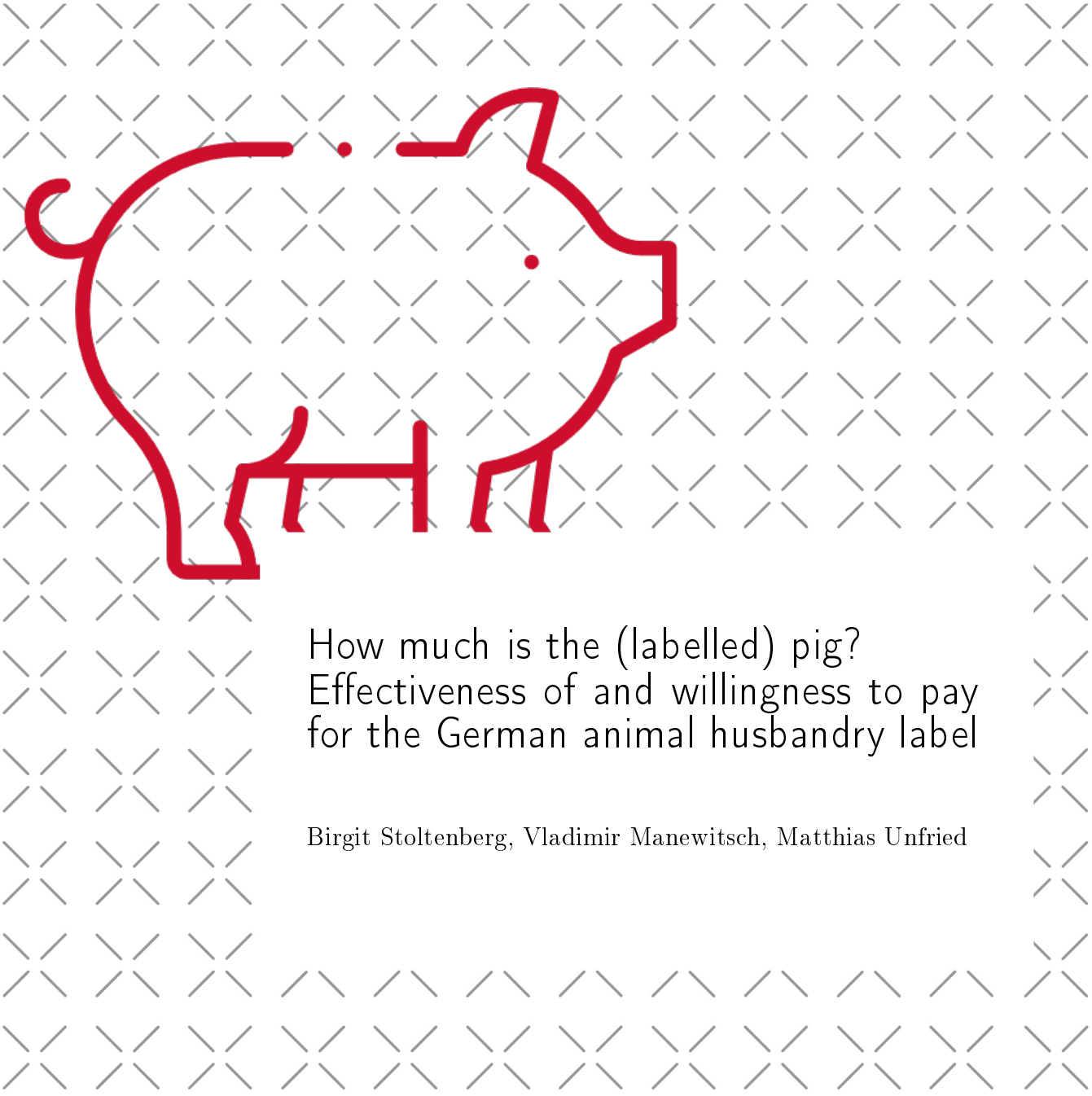
NIM Nuremberg Institute for Market Decisions



How much is the (labelled) pig? Effectiveness of and willingness to pay for the German animal husbandry label

Birgit Stoltenberg, Vladimir Manewitsch, Matthias Unfried

NIM WORKING PAPER SERIES > 8 > 2023

Publisher: Nürnberg Institut für Marktentscheidungen e.V. / *Founder of GfK*

How much is the (labelled) pig? Effectiveness of and willingness to pay for the German animal husbandry label

Birgit Stoltenberg*† Vladimir Manewitsch† Matthias Unfried†

Abstract— *The German Minister of Food and Agriculture has initiated a mandatory state label for animal husbandry, revealing the livestock production conditions. This study investigates consumer preferences and their willingness to pay price premiums for pork depending on the different husbandry conditions indicated by this label. We tested the originally planned label, as well as a modification with additional information on the available living space for the animal, a modification with additional interpretive information (i.e., with colours indicating how a product scores on animal husbandry standards), and a combination of both modifications. Data were obtained from a representative online discrete choice experiment ($N = 2,015$) in Germany and analysed using random parameter logit modelling. Our results show that the label proposed by the government is effective in increasing the willingness to pay for meat produced with higher standards, except 'organic' level. Furthermore, we show that indicating a ranking of husbandry conditions could further increase the willingness to pay for meat produced with higher husbandry standards.*

Keywords— *Animal welfare; Attitude–behaviour gap; Credence attribute; Discrete choice experiment (DCE); Multi-level label; Organic; Purchase decision; Sustainable consumption; Willingness to pay;*

1 Introduction

Meat production and consumption have faced increasing criticism in recent years. Concerns focus on the negative ecological impacts, such as disproportionate land use, greenhouse gas emissions, water extraction, and biodiversity loss (Godfray et al., 2018; Aleksandrowicz et al., 2016; Steinfeld et al., 2006), as well as the conditions under which animals are raised for meat production (Busch et al., 2018). This debate gained further momentum when the German Minister of Food and Agriculture initiated the mandatory state labelling of livestock conditions in 2022 (Bundesministeriums für Ernährung und Landwirtschaft (BMEL), 2022b, 2022c).

Consumer preferences also seem to reflect a growing interest in the welfare of animals raised for meat production. According to a consumer survey conducted by the German Ministry of Food and Agriculture, approximately 80% of German consumers indicated that livestock production conditions matter when making meat purchases (BMEL, 2022a). However, market data paint a different picture, revealing a significant gap between consumer attitudes and purchasing behaviour. For example, organic pork, involving a production method with rather strict husbandry standards, accounts for only 1.78% of the German pork market in 2022 (Bund Ökologische Lebensmittelwirtschaft e. V. (BÖLW), 2023).

Several factors contribute to this attitude–behaviour gap. On the one hand, there is evidence that consumers are willing to pay a premium for meat that is produced in an animal welfare–friendly fashion (Janssen et al., 2016) or using non-conventional methods, such as organically raised or grass-fed beef (Risius & Hamm, 2017; Xue et al., 2010). On the other hand, price remains a key deterrent for consumers when it comes to purchasing non-conventional food (Aschemann-Witzel & Niebuhr Aagaard, 2014). Additionally, quality considerations play a significant role in consumer decision making, as identified by Buder et al. (2014).

Assessing the quality of a product—for instance, the husbandry conditions of livestock—can be challenging. Consumers generally evaluate product quality based on three different types of attributes: observable attributes before purchase, so-called search attributes (Stigler, 1961),

*Corresponding author, birgit.stoltenberg@nim.org

†Nuremberg Institute for Market Decisions, Nuremberg, Germany

Citation: Stoltenberg, B., Manewitsch, V., & Unfried, M. (2023). How much is the (labelled) pig? Effectiveness of and willingness to pay for the German animal husbandry label, *NIM Working Paper Series, 8.*

e.g., price and colour; attributes experienced after purchase, named experience attributes (Nelson, 1970), e.g., the taste of the product and the convenience of use; and attributes that are not observable at all, i.e., credence attributes (Darby & Karni, 1973), such as livestock husbandry conditions (Grunert et al., 2004; Bredahl et al., 1998). For credence attributes, consumers have to rely on information provided by manufacturers, retailers, or other external parties. If there is uncertainty or doubt about the quality of the information, consumers are less willing to pay higher prices, and suppliers lack incentives to offer higher–quality products (Akerlof, 1970).

Information on credence attributes is necessary to bridge the attitude–behaviour gap and enable consumers to make informed choices according to their preferences. Product labels can serve as a valuable source of credence information. Labels indicating the livestock husbandry conditions, for example, can provide consumers with an indication of a product's quality, reduce prepurchase evaluation costs, and facilitate purchasing decisions based on personal preferences (Thøgersen et al., 2012). Furthermore, labels remind consumers of sustainable and ethical issues and therefore focus consumers' limited attention at the point of sale (Peschel et al., 2019).

Ikonen et al. (2020) classified (nutrition) labels into reductive labels and interpretive labels. While reductive labels provide information without offering any interpretation of this information, interpretive labels offer greater evaluation of presented information and can be further categorized depending on the aggregation of information (Newman et al., 2018; Talati et al., 2017). There are binary (or simple) labels with a positive or negative message and multi-level labels with the possibility of communicating more information. The latter ones might be a more appropriate way to communicate product attributes like animal husbandry conditions (Tonsor & Wolf, 2011; Torma & Thøgersen, 2021; Weinrich & Spiller, 2016).

In the food market, there are various multi-level labelling systems indicating the extent to which a product fulfils a certain criterion. In the EU egg market, for instance, numbers indicate the type of poultry keeping (Janssen et al., 2016). However, these labels provide no information to relate this to alternative standards of livestock farming; thus, an uninformed consumer might be unable to classify the product within the range of possible husbandry conditions.

Product labels play a crucial role in conveying information to consumers, but their effectiveness depends on several key conditions, as highlighted in the literature. First and foremost, for labels to have an impact on purchasing decisions, consumers need to be aware of them. Numerous studies (Thøgersen, 2000; Leire & Thidell, 2005; Aertsens et al., 2011; Miller & Cassady, 2015) emphasize the importance of label awareness, as consumers must be informed about the label's existence and content.

In addition to awareness, the perceivability of product labels is another critical factor. Research by Thøgersen (2000), De Bauw et al. (2021), Packer et al. (2021), and Stoltenberg et al. (2022) underlines the significance of a label to be understood. Consumers should be able to easily interpret and comprehend the information provided on the label. If labels are confusing or complex, their effectiveness diminishes, as consumers may struggle to make sense of the information presented.

Trust also plays a crucial role in the effectiveness of product labels. Sirieix et al. (2013) and Tonkin et al. (2015) highlight the importance of trust in the label publisher. Consumers are more likely to place trust in labels that are issued by reputable sources, such as governmental organizations. Majer et al. (2022) summarise that governmental labels tend to be more reliable and trustworthy than non-governmental ones. Moreover, trust in a product label can be supported by the use of standardized, sound, and verified criteria, including adherence to legal principles.

By meeting these conditions, product labels can effectively inform and guide consumers in making informed purchasing decisions. An example of a label that is verifiable and gives consumers the right amount of information is the Nutri-Score. It indicates a food product's nutritional value on a multi-level scale using colours and letters, ranging from the highest nutritional value, A (green), to the lowest, E (red). The introduction of the Nutri-Score label effectively helped consumers to understand nutritional values (De Bauw et al., 2021); the label seems to be a suitable way for consumers to classify products as healthy or unhealthy (Packer et al., 2021).

A slightly different approach was used in Germany, where retail and industry introduced a voluntary label in 2019 to classify livestock husbandry conditions. This label, an interpretive multi-level label, distinguishes between four levels of animal husbandry, ranging from level 1, housing (the legal minimum standard), to level 4, organic. In 2022, the German Minister of Food and Agriculture initiated the launch of a mandatory state label that indicates livestock husbandry conditions across five different levels. The suggested label is designed as a reductive monochrome multi-level label (cf. Figure 1). Both animal husbandry labels aim to provide consumers with information about husbandry conditions. However, both labels currently employ only text-based information, lacking detailed information about the specific husbandry conditions of the respective husbandry level. Consequently, the distinction between different levels of husbandry conditions is not immediately apparent, and this lack of information may result in consumers falling back on heuristics and implicitly making assumptions about the relationship between the levels (Stoltenberg et al., 2022). This could lead to a misperception of animal husbandry conditions and biased consumption patterns, as purchase decisions may not reflect consumers' actual preferences and compromise the label's supposed impact in supporting sustainable consumption.

Although the ideal design of a label is still unclear (Torma & Thøgersen, 2021; Donato & Adigüzel, 2022), this effect might be avoided when a multi-level label

provides more detailed information about the relation between the label levels and their underlying basis or implications. This could be done by either:

a) Adding an informational layer with additional animal husbandry–specific information regarding the main characteristic of the husbandry conditions (e.g., available space per animal) and making the difference between the label levels more salient and accessible directly on the label rather than via a QR code.

b) Adding an interpretive layer, evaluating how good a product scores on this aspect. This additional information (e.g., a colour scale for the different levels, similar to the Nutri-Score colours) facilitates consumers' understanding of the message (Roberto et al., 2012).

Both additional layers of information can help the consumer to assess the difference between the husbandry conditions, making it easier for them to make purchase decisions according to their preferences. At least the colour scale approach was shown to be highly effective (Song et al., 2021).
In this paper, we present a discrete choice experiment (DCE) in which the design of the planned German mandatory state label for animal husbandry is experimentally varied. We test whether such labels can fulfil their purpose of providing meaningful information about livestock husbandry conditions, or if there is a need to enhance the salience of these conditions and the associated differences. Specifically, we aim to answer the following research questions (RQ):

RQ1: Effectiveness of the planned mandatory state label for animal husbandry, a reductive monochrome multilevel label: Do higher animal husbandry standards implied by the label increase the willingness to pay?

RQ2: Increased effectiveness of the planned mandatory state label for animal husbandry with additional informational layer: Does additional information about available space per animal increase the willingness to pay?

RQ3: Increased effectiveness of the planned mandatory state label for animal husbandry as interpretive multilevel label by giving red to green colors to different levels: Does an additional interpretive layer increase the willingness to pay?

RQ4: Increased effectiveness of the planned mandatory state label for animal husbandry with both additional layers, the information about the available space per animal and the interpretive colour scale: Does the combination of both additional layers increase the willingness to pay?

The remainder of this paper is organised as follows. Section 2 introduces the method, including the experimental design, the sample, and the estimation approach. Section 3 presents the results. In section 4, we discuss the results and the conclusion of the study.

2 Method

2.1 Experimental design

We conducted an experimental study using a discrete choice experiment (DCE). DCE is a choice-based survey method that elicits people's preferences for different options in realistic choice scenarios instead of asking them directly about their preferences. We chose this method because people tend to give socially desirable answers when asked directly. This method furthermore allows to assess the willingness to pay and the influence of certain product attributes on the willingness to pay. To get a more realistic picture of the preferences, we used a special type of DCE, the dual response DCE.
To this end, we presented hypothetical purchase scenarios where participants had to choose one out of three products in ten rounds. The products were described by different attributes: price, packaging, region of origin, and animal husbandry conditions indicated by the planned mandatory state label (the basic version of it is shown in Figure 1). A detailed description of the DCE design is given later in this section.

Figure 1

Basic version of the planned mandatory state label for animal husbandry

Note. This was the planned version until March 2023, when 'Auslauf/Freiland' was changed to 'Auslauf/Weide'.
English translation: Bio = Organic; Auslauf/Freiland = Outdoor runs/free-range; Frischluftstall =Indoor with fresh air; Stall+Platz = Indoor+space; Stall = Indoor housing.

To analyse the influence of animal husbandry conditions on label effectiveness and willingness to pay, as well as the influence of additional information on the label, we experimentally manipulated the design of the animal husbandry label and added information on one criterion for animal husbandry. We chose available space per animal because it seemed to be one of the most important criteria for the perceived quality of animal husbandry (Janssen et al., 2016), which was confirmed in a short pre-study.
In the control group ($G_{original}$), the original label was used. In the first experimental group (G_{space}), we added a number indicating the area of living space in square meters the animal has at the respective husbandry level and a visualisation of the living space relative to the size of an average animal. In the second experimental group (G_{colour}) with an additional interpretive layer, we coloured the label—similar to the Nutri-Score label or the European Energy Efficiency label—to add an interpret-

Table 1

Experimentally varied design of the animal husbandry label

	$G_{original}$	G_{space}	G_{colour}	$G_{space+colour}$
Indoor housing				
Indoor+ space				
Indoor with fresh air				
Outdoor runs\ free-range				
Organic				

Note. Example for the stimuli used: control group (original), additional informational layer (space), additional interpretive layer (colour), and combination of both layers (space+colour).

ive information for the animal husbandry, though the order of the levels on the label was left unchanged. The colour scale includes a letter from A (green) to E (red) as with the aforementioned labels. In the third experimental group ($G_{space+colour}$), we combined both layers, adding information about living space per animal as well as a colour scale. Participants were randomly assigned to one of the four groups. The labels used in the study are depicted in Table 1.[1]

Before and after the DCE, participants were asked to answer a questionnaire. The pre-questionnaire included only two sociodemographic questions (age and gender) for screening and quotation and questions on meat purchase behaviour (frequency of purchase and consumption, preferred retailer, and out-of-home consumption). In the post-questionnaire, participants were asked detailed sociodemographic questions, as well as their perception, knowledge, and preferences about the animal husbandry label. Finally, they were asked about their general usage and awareness of (sustainability) labels.

2.2 Sample

The survey was conducted with $N = 2,015$ participants in Germany. The sample in each group was representative with respect to age and gender for the German online population. Participants were recruited by a market research company from their commercial market research panel and screened for pork consumption. The study was conducted in February and March 2023 using the survey software *QuestionPro*.

The average age of the participants was 49.47 years ($SD = 16.73$). Approximately 50.82% of the participants identified as female, and 48.44% identified as male; 0.74% of the participants answered 'diverse'. The respondents received a participation fee of €2.00, and it took the participants 13.5 minutes ($SD = 17.06$) on average to complete the study. Table 2 provides a descriptive summary of the sample.

2.3 Design of the discrete choice task

Corresponding to our research question, we chose pork cutlet as the product for the choice tasks. The attributes describing the product were chosen according to the literature: the husbandry condition indicated by the planned mandatory state label, the price, the packaging, and the region of origin. The price is later used to calculate willingness to pay and ranges from €2.19/250g to €6.29/250g. The price levels were selected according to the average prices of conventional and organic pork cutlets sold in conventional and organic German supermarkets. Thus, we considered the following four attributes with their corresponding levels (see Table 3).

In the questionnaire, the respondents were provided information about the procedure as well as detailed descriptions of the product attributes and levels before the DCE started.

To create the description of the products for the DCE design (the so-called product cards), we used the R package 'DoE.base'[2] to access orthogonal arrays (Grömping,

[1] The drafts of the German Ministry of Food and Agriculture also contain a label version with a coloured background. For the sake of completeness, we tested this alternative version with an additional group.

[2] Version 1.2-1

Table 2

Descriptive summary statistics for the experimental groups

	$G_{original}$	G_{space}	G_{colour}	$G_{space+colour}$
N	503	504	503	505
Gender (ratio female in %)	51.20	51.20	51.20	51.20
Age	49.48	49.67	49.51	49.22
Education (4-point scale)[a]	2.29	2.29	2.29	2.27
Household size	1.97	1.93	1.96	1.99
Financial situation (5-point scale)[b]	2.87	2.75	2.87	2.83
Residence size (5-point scale)[c]	3.00	2.92	2.90	2.90

[a] 1 = (Still) without vocational qualification; 2 = Apprenticeship/vocational training/technical college degree; 3 = Degree; 4 = Doctorate
[b] 1 = I do not have to restrict myself in any way; 2 = I am well provided for and can afford some things; 3 = On the whole I get by; 4 = I can just about make ends meet; 5 = I do not have enough money at all
[c] 1 = Under 5,000 inhabitants; 2 = 5,000 − 20,000 inhabitants; 3 = 20,001 − 100,000 inhabitants; 4 = 100,001 − 500,000 inhabitants; 5 = Over 500,000 inhabitants

Table 3

Attributes and levels used in the hypothetical choice experiment

Attribute	Level	
Type of Husbandry (ToH)	1	**Indoor housing**[a]
	2	Indoor+space[b]
	3	Indoor with fresh air[c]
	4	Outdoor runs/free-range[d]
	5	Organic[e]
Price (€/250g)		€2.19; €2.69; €3.19; €3.69; €4.29; €4.89; €5.59; €6.29
Packaging (Pac)	1	**Packaged from the refrigerated section**[f]
	2	Fresh off the meat counter[g], later referred to with 'fresh'
Region of Origin (RoO)	1	**From Germany but outside your region**[h]
	2	From your region within a radius of 100km[i], later referred to with 'region'

Note. Levels in bold are reference levels in the model estimation. The original German text appears in footnotes.
[a] Stall [b] Stall+Platz [c] Frischluftstall [d] Auslauf/Freiland [e] Bio
[f] Verpackt aus dem Kühlregal [g] Frisch von der Fleischtheke
[h] Aus Deutschland, aber außerhalb Ihrer Region
[i] Aus Ihrer Region im Umkreis von100 km

2018). This resulted in 80 cards. To exclude unrealistic options, we restricted the orthogonal design and excluded combinations of the lowest husbandry conditions with the highest prices and the highest husbandry conditions with the lowest prices. These restrictions resulted in 68 possible combinations. To ensure level balance for each attribute, we further deleted eight combinations, resulting in 60 remaining cards. The remaining 60 cards have a D-efficiency of 28.36 (Street & Burgess, 2007).

The 60 cards were assigned randomly to 20 choice tasks, followed by a manual step in case one choice task contained a dominant option and also to improve utility balance (Huber & Zwerina, 1996). This procedure resulted in choice tasks with realistic purchase options. The blocking technique was applied to further reduce the number of choice sets for each participant. One of two blocks with ten choice tasks each was randomly presented to each participant. In summary, our DCE design was constructed to ensure level balance while approaching orthogonality, minimal overlap, and utility balance (Huber & Zwerina, 1996).

The cards in each choice task described the product using the four attributes described above. Thus, each respondent was exposed to ten choice tasks, each consisting of three cards, followed by a so-called dual response none question. Respondents first had to choose the most preferred alternative out of the three options presented ('forced choice'). After expressing their preference, respondents were asked whether they would actually buy the product in a supermarket, so the second question includes a no-choice option ('free choice'). This combination of a forced choice followed by a free choice is called dual response DCE (Brazell et al., 2006; Diener et al., 2006; Dhar & Simonson, 2003; Heidel et al., 2021). Through dual response DCE the available dataset contains more choice information, at least in situations where the no-choice option is likely to be chosen (Brazell et al., 2006).

Alternatives in each choice set were presented in a randomised order and sequence. An example of the choice tasks is depicted in Figure 2.

In this study we only analyse data from the forced choice tasks (question 7a), not the dual response free choice task (question 7b).

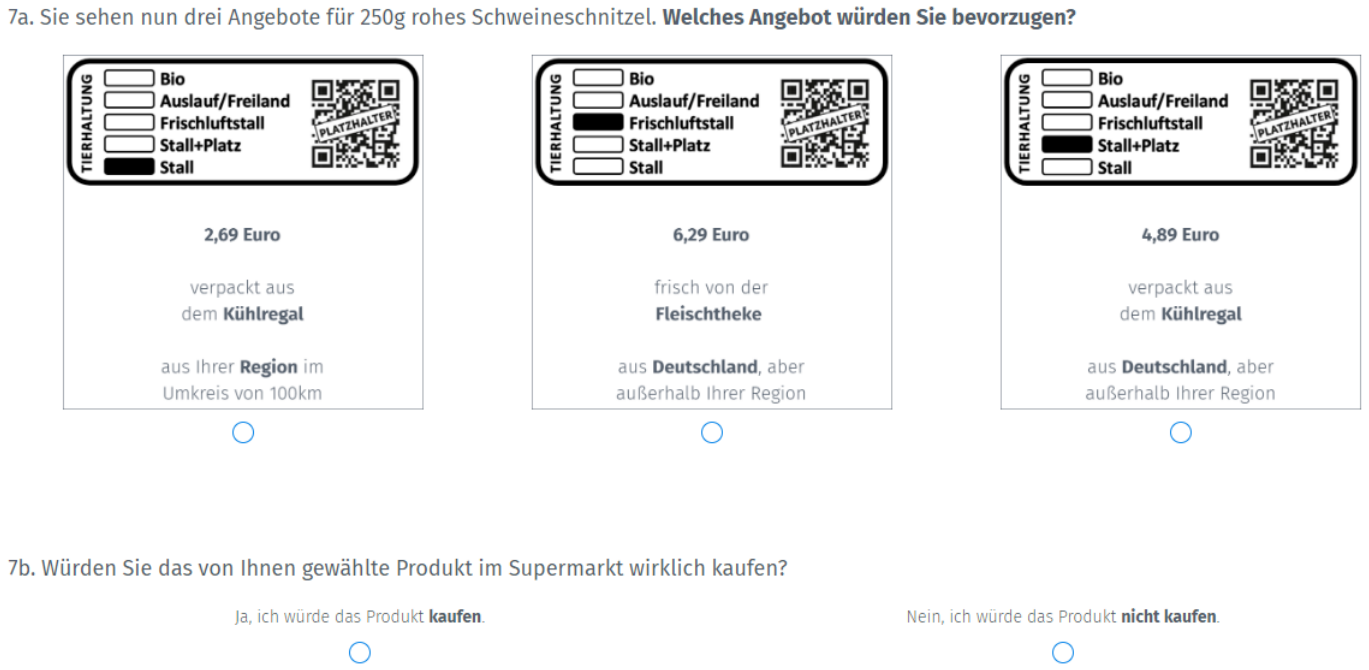
2.4 Estimating willingness to pay

To assess the willingness to pay for the designed attribute levels and the corresponding differences between the experimental groups, we used the standard random utility model according to McFadden (1974). The model assumes that the alternative with the highest total utility is chosen in a choice task. Furthermore, it models the utility as a linear regression function of the presented attribute levels and an error term representing factors not accounted for.

To control for potential significant preference heterogeneity with respect to animal welfare aspects, we utilised the random parameter or mixed logit formulation of the model, which allows the parameters of the linear utility

Figure 2

Sample choice task for the control group



Note. English Translation: 7a. Sie sehen nun drei Angebote für 250g rohes Schweineschnitzel. **Welches Angebot würden Sie bevorzugen?** = 7a. You now see three offers for 250g of raw pork cutlet. **Which offer would you prefer?**
7b. Würden Sie das von Ihnen gewählte Produkt im Supermarkt wirklich kaufen? Ja, ich würde das Produkt **kaufen**.—Nein, ich würde das Produkt **nicht kaufen**. = 7b. Would you actually buy the product you have chosen in the supermarket? Yes, I would **buy** the product.—No, I would **not buy** the product.

function to vary across respondents (Hensher & Greene, 2003). Thus, the model is able to reflect the panel structure of the dataset with repeated measurements for each respondent. It also allows for the relaxing of the IIA (independence of irrelevant alternatives) assumption (Train, 2009) of the starting model, which is unlikely to hold for different animal husbandry label levels. We imposed that the random parameters of all attributes except price are multivariate and normally distributed across the respondents. For the (negative) price parameter, we assumed a log-normal distribution to ensure its identification (Daly et al., 2012). Hence, our model assumes that preference heterogeneity is reflected by (individual) shifts in the mean of that distribution. As we did not apply any further restrictions on the total covariance matrix of the random parameters, all utility parameters are allowed to vary across respondents retaining their logical dependencies (e.g., a higher preference for higher husbandry levels logically goes along with a lower preference for lower husbandry levels) as well as scale heterogeneity related sources of dependences (Hess & Train, 2017).
Based on considerations from Train and Weeks (2005) and Scarpa et al. (2008), we tested both formulations of the model, i.e., in the preference space as well as in the willingness to pay space. As the estimation directly in the willingness to pay space showed lower fit[3] and less reasonable distributions of the estimates, we postulated our estimation model in the preference space and derived willingness to pay estimates for all attribute levels by dividing the corresponding utility parameters by the price parameter.
Finally, to evaluate the differences between the experimental groups, we combined their data into a joint dataset and postulated a combined model equation (1) by dummy-coded attribute levels (except price):

$$\begin{aligned} U_i = \beta_{i2}ToH_2 + \beta_{i3}ToH_3 + \beta_{i4}ToH_4 + \beta_{i5}ToH_5 + \\ + \gamma_i Price + \beta_{ip}Pac_{fresh} + \beta_{ir}RoO_{region} + \\ + \sum_T G_i^T(\beta_{i2}^T ToH_2 + \beta_{i3}^T ToH_3 + \beta_{i4}^T ToH_4 + \beta_{i5}^T ToH_5) \end{aligned} \quad (1)$$

where U_i is the explained part of utility for respondent i (omitting indexing of alternatives and choice tasks to reduce notation clutter). The variables ToH_2, ToH_3, oH_4, ToH_5 are dummy-coded levels for the type of husbandry, the levels as reported in Table 3 (2 corresponds to 'indoor+space', 3 to 'indoor with fresh air', 4 to 'outdoor runs/free-range', and 5 to 'organic'). The variables Pac_{fresh} and RoO_{region} are dummy-coded levels for fresh unpacked and regional products correspondingly. $\beta_{i\_}$ denotes the individual parameters for alle attributes besides price (type of husbandry, packaging and region of origin), and γ_i the individual price parameter. G_i^T is a treatment dummy for the experimental groups *space*, *colour*, and *space+colour*. As usual, one dummy is omitted for each

[3] All model quality criteria (log-likelihood, AIC, and BIC) showed better fit for the model estimated in preference space compared to the one estimated in willingness to pay space.

attribute to ensure identification so that all β coefficients are interpreted as marginal effects relative to the corresponding omitted levels (i.e., ToH_1, Pac_1 and RoO_1), cf. Table 3.

The parameters β_{i2}, β_{i3}, β_{i4}, β_{i5} therefore represent marginal utilities in the control group for the husbandry levels 2, 3, 4, and 5, relative to the utility of level 1. The experimental group-specific coefficients β^T_{i2}, β^T_{i3}, β^T_{i4}, β^T_{i5} are the marginal treatment effects of the husbandry levels 2, 3, 4, and 5 for each experimental group $T \in \{space, colour, space + colour\}$, relative to the corresponding marginal utilities in the control group. Thus, the model assumes that there is no interaction between treatment and attributes for price, packaging, and region of origin, as our experimental manipulations only affect the animal husbandry label.

The model is estimated by the simulated maximum likelihood procedure (Train, 2009) based on 2,000 Halton draws. The full log-likelihood of the model is maximised via the Berndt–Hall–Hall–Hausman algorithm by the R package 'gmnl'[4] (Sarrias & Daziano, 2017).

The estimate of the willingness to pay for a certain product attribute level is calculated by dividing $\hat{\beta}$ (mean estimates for the β parameters) by $\hat{\gamma}$ (the mean estimate of the price parameter γ). For the calculation of the standard errors of the resulting willingness to pay estimates, the delta method of R package 'msm'[5] (Jackson, 2011) is used.

Table 4

Results of the mixed logit willingness to pay estimates for the control group and marginal treatment effects

	$G_{original}$	ΔG_{space}	ΔG_{colour}	$\Delta G_{space+colour}$
ToH_2	1.02***	0.01	0.37*	0.04
	(0.094)	(0.136)	(0.146)	(0.158)
ToH_3	1.99***	-0.21	0.57***	0.53***
	(0.090)	(0.137)	(0.131)	(0.141)
ToH_4	3.05***	-0.40**	0.66***	0.43**
	(0.090)	(0.139)	(0.132)	(0.143)
ToH_5	2.12***	-0.09	0.93***	0.86***
	(0.123)	(0.179)	(0.181)	(0.188)
Pac_{fresh}	0.27***			
	(0.026)			
RoO_{region}	0.34***			
	(0.027)			

Note. Significance codes: ***p <0.001, **p <0.01, *p<0.05, . p<0.1
Standard errors in parentheses.

[4] Version 1.1.3.2

[5] Version 1.7

3 Results

Table 4 shows the estimation results for the postulated model, i.e., for the $\hat{\beta}/\hat{\gamma}$ quotients of parameters in equation (1). The estimates correspond to the mean marginal willingness to pay across respondents for the respective attribute levels; their standard errors are given in parentheses.

Column $G_{original}$ represents the marginal willingness to pay in the control group for the animal husbandry conditions 2, 3, 4, and 5 relative to level 1, as well as the marginal willingness to pay for fresh and regional pork cutlets relative to the packaged and national origin products, respectively. Columns ΔG_{space}, ΔG_{colour}, and $\Delta G_{space+colour}$ show the (additional) average effects of our experimental manipulations compared to the corresponding marginal willingness to pay values in the control group.

The estimates in Table 4 can be interpreted as additional willingness to pay relative to the (omitted) reference levels or to the control group for the respective label level. For example, respondents in the control group are willing to pay €1.02 more for a 250g pork cutlet with the animal husbandry condition 2 'indoor+space' compared to the same piece of meat with the animal husbandry condition 1 'indoor housing'. In the case of the interpretive modified label with additional colour scale, they will pay €0.66 more for a pork cutlet with the label level 4 'outdoor runs/free-range' compared to the original label level 4 from $G_{original}$. We can even compare label level 1 from the control group with, e.g., label level 3 'indoor with fresh air' from $\Delta G_{space+colour}$, by adding the monetary effects: €1.99 + €0.53 = €2.52 will be paid, on average, additionally for a modified label with both additional layers with a higher husbandry standard.

The estimates for the control group in column $G_{original}$ show that the original label design induces significantly higher willingness to pay for all husbandry label levels 2, 3, 4, and 5 compared to level 1. It increases with the label levels 2, 3, and 4 but drops from levels 4 to 5. This seemingly surprising finding that the willingness to pay for the organic level 5 is not the highest will be discussed in the next section. Column $G_{original}$ also shows a significantly higher willingness to pay for fresh packaged meat as well as for regional meat. In summary, these results imply that animal husbandry is an important factor for consumers. However, our finding for organic level 5 suggests that organic production does not induce a higher willingness to pay compared to outdoor runs/free-range level 4.

The results in column ΔG_{space} indicate a decrease in the willingness to pay between levels 4 and 1 for the label with additional numeric and visual information about the available space for the animal in m^2. For levels 2, 3, and 5, the corresponding changes in willingness to pay are not statistically significant. These results suggest a decrease in willingness to pay for level 4 relative to the other levels. In summary, the additional informational layer appears to have only a weak and negative effect on the willingness to pay.

For the label with an additional interpretive layer, i.e., the colour-coded levels including a letter, the results in column ΔG_{colour} show an increase in willingness to pay for all levels relative to level 1. These marginal changes increase for higher levels of the manipulated label. These results indicate a strong positive effect of interpretation–aiding colour scale on the willingness to pay.
When both additional layers of information are combined in one label, as shown in column $\Delta G_{space+colour}$, the results suggest an approximately additive pattern of their impacts without strong interaction. The statistically significant increase in willingness to pay for levels 3, 4, and 5 of the animal husbandry label is lower than the corresponding increase for the interpretation–aiding colour scale information only in column ΔG_{colour} but higher than the changes for the information about available space only.
In total, the impact of the colour scale seems to be stronger than that of the available space per animal, indicating that respondents pay more attention to the (familiar) colour-coded evaluation scheme than to (novel) information cues.

4 Discussion, limitations, and conclusion

4.1 Summary and discussion

The German Minister of Food and Agriculture suggested a mandatory state labelling of livestock conditions using a reductive monochrome multi-level label. In this study, we tested the effectiveness of this label with respect to willingness to pay and whether design adjustments could improve the effectiveness.
Our results confirm the findings of the existing literature on multi-level labels—namely, that these labels are an appropriate way to communicate product attributes like animal husbandry conditions (Tonsor & Wolf, 2011; Torma & Thøgersen, 2021; Weinrich & Spiller, 2016). However, the design and the information that is presented on the label are important. We found that the proposed label is effective in the sense that consumers are willing to pay price premiums for meat that is produced under higher animal husbandry standards (label levels $2-5$) compared to meat that is produced with the legal minimum standard (label level 1). However, the willingness to pay is not monotonically increasing with the label level, as the willingness to pay for the highest husbandry standard (label level 5 'organic') is lower than the one for label level 4.
So far, we can only speculate about the reasons for this result. In a pre-study, we asked the respondents to rank the five categories based on the wording without showing the label to them.[6] We found that consumers seem to have difficulties in ordering the label levels according to their names, especially to classify the level 'organic' hierarchically. Levels 1 to 4 could clearly be ranked according to animal husbandry conditions, but 'organic' was ranked worst, second best and best in almost equal proportions.
In our study, we asked the experimental groups without additional information on space availability ($G_{original}$ and G_{colour}) to estimate the total space (indoor and outdoor) for each animal. Although the participants estimated the difference in space per animal for levels 1 and 2 rather well, they tend to underestimate the difference from level 2 to 3 and overestimate the difference between levels 3 and 4. The difference in husbandry conditions between levels 4 and 5 where the increase in space is highest was drastically underestimated by the participants. On average, participants assume less than 10% space improvement being an 'organic' versus an 'outdoor runs/free range' pig.
One reason could be that 'organic' is perceived as an indicator for healthy food rather than for animal husbandry conditions (Aertsens et al., 2011); since the names of the other levels are clearly related to husbandry conditions, the word 'organic' leaves the specific conditions unclear. Another reason could be that German respondents are already used to the existing voluntary animal husbandry label with only four levels, in which label level 4, named 'premium', has a broader focus and includes organic standards but is not restricted to them. Nevertheless, further explorative analyses showed that there might be subgroups—e.g., younger consumers and consumers who are financially better situated—whose willingness to pay increases steadily and is the highest for organic meat.
Regarding the experimental manipulation of the husbandry level, we did not find (positive) effects of adding living space information. The numeric and visual information regarding the label's underlying criteria on square meters of animal living space does not increase the willingness to pay. The results are not significant and, thus, cannot be interpreted. A significant exception is label level 4, which creates a lower willingness to pay than the corresponding label level of the control group. This effect of label level 4 is probably related to the drop in the indoor living space ($1m^2$) compared to that in the lower level 3 ($1.3m^2$).
The additional interpretive layer (colour codes and letters from red (E) to green (A) indicating a ranking) is able to increase the willingness to pay for all label levels. All results are positive and significant and the increase in the willingness to pay is even monotonic. It ranges from €0.37 for level 2 to €0.93 for 'organic'.
Adding both, information about living space and a colour scale made the label more effective than the original version. However, compared to the label with colour scale only, the combined information resulted in a slightly lower willingness to pay.

[6] All questions visually showing the proposed state label for animal husbandry were asked after this ranking question.

4.2 Limitations and avenues for further research

As our study was not incentivised, the choices in the DCE were hypothetical for the respondents. This limitation belongs to the chosen method (Haghani et al., 2021; Schulze et al., 2021; Sonntag et al., 2023). Further experimental insights observing real purchasing behaviour can yield further insights.
We presented the DCE tasks in traditional layout, i.e., the pork cutlets were described with a list of attributes. On a photographic image of the products including all the information directly on the package (for packaged products) or in realistic presentation (for fresh products), the attributes could be displayed in real size ratio, which also could add to realistic results.
Our study was conducted in a period of high inflation in Germany, which means that the respondents may have been especially sensitive to prices.
The study was conducted on pork meat because this type of meat will be the first to use this form of mandatory labelling. Different animals require different animal welfare standards, so our results may not be fully generalizable to other types of meat.
Our study showed an interesting and unexpected effect: why is organic meat valued less compared to meat produced under the 'outdoor runs/free-range' condition. So far, there is no explanation on this effect. At least the requirements for organic meat are much higher than those for 'outdoor runs/free-range' meat and organic conditions include even more outdoor space available for each animal (at least twice as much). But as long as the majority of the consumers is not aware of details on organic husbandry conditions, this could result in lower willingness to pay. So, further research is necessary on the label level 'organic'.
Finally, research on providing detailed information about husbandry standards via QR codes can be an interesting avenue to pursue.

4.3 Conclusion

Our study shows that the proposed label is effective, consumers are willing to pay price premiums for meat that is produced under higher animal husbandry standards. Consumers willing to pay more are an important enabling factor for the transformation towards higher animal husbandry standards. Then, farmers who convert their barns and have higher costs for animal husbandry could expect to cover these costs by additional income.
However, the willingness to pay for the organic level is not highest, although the husbandry standards are most strict in this category. Here, the conditions farmers have to fulfil being certified as organic meat producer have to be more communicated to reach the majority of consumers, as only subgroups valued the highest animal husbandry standards by highest willingness to pay.
The design and the information that is presented on the label are important. We were able to show that colour codes indicating an ordered ranking can significantly influence purchase decisions and increase willingness to pay for meat produced with higher husbandry standards. These findings suggest that the proposed label could communicate the differences between the animal husbandry conditions even better and more effectively—and would thus enable consumers to shop more in line with their preferences and thus have a more effective steering impact on animal husbandry in Germany.

References

Aertsens, J., Mondelaers, K., Verbeke, W., Buysse, J. & Van Huylenbroeck, G. (2011). The influence of subjective and objective knowledge on attitude, motivations and consumption of organic food. *British Food Journal*, *113*(11), 1353-1378.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, *84*(3), 488-500.

Aleksandrowicz, L., Green, R., Joy, E. J., Smith, P. & Haines, A. (2016). The impacts of dietary change on greenhouse gas emissions, land use, water use, and health: A systematic review. *PLOS One*, *11*(11).

Aschemann-Witzel, J. & Niebuhr Aagaard, E. M. (2014). Elaborating on the attitude–behaviour gap regarding organic products: Young Danish consumers and in-store food choice. *International Journal of Consumer Studies*, *38*(5), 550-558.

Brazell, J. D., Diener, C. G., Karniouchina, E., Moore, W. L., Séverin, V. & Uldry, P. F. (2006). The no-choice option and dual response choice designs. *Marketing Letters*, *17*, 255-268.

Bredahl, L., Grunert, K. G. & Fertin, C. (1998). Relating consumer perceptions of pork quality to physical product characteristics. *Food Quality and Preference*, *9*(4), 273-281.

Buder, F., Feldmann, C. & Hamm, U. (2014). Why regular buyers of organic food still buy many conventional products: Product-specific purchase barriers for organic food consumers. *British Food Journal*, *116*(3), 390-404.

Bund Ökologische Lebensmittelwirtschaft e. V. (BÖLW). (2023). *Branchenreport 2023 – Ökologische Lebensmittelwirtschaft.* `https://www.boelw.de/fileadmin/user_upload/Dokumente/Zahlen_und_Fakten/Broschuere_2023/BOELW_Branchenreport2023.pdf`. (Accessed: 2023-05-02)

Bundesministeriums für Ernährung und Landwirtschaft. (2022a). *Deutschland, wie es isst: Der BMEL-Ernährungsreport 2022.* `https://www.bmel.de/SharedDocs/Downloads/DE/Broschueren/ernaehrungsreport-2022.pdf?__blob=publicationFile&v=9`. (Accessed: 2023-06-06)

Bundesministeriums für Ernährung und Landwirtschaft. (2022b). *Gesetzentwurf der Bundesregierung: Gesetz zur Kennzeichnung von Lebensmit-*

teln mit der Haltungsform der Tiere, von denen die Lebensmittel gewonnen wurden. https://www.bmel.de/SharedDocs/Downloads/DE/Glaeserne-Gesetze/Kabinettfassung/tierhaltungskennzeichnungsgesetz-kabinett.pdf?__blob=publicationFile&v=6. (Accessed: 2023-06-06)

Bundesministeriums für Ernährung und Landwirtschaft. (2022c). *Zukunftsfeste Tierhaltung: Eckpunkte zur Einführung einer verpflichtenden staatlichen Tierhaltungskennzeichnung.* https://www.bmel.de/SharedDocs/Downloads/DE/_Tiere/Tierschutz/eckpunkte-tierhaltungskennzeichnung.pdf?__blob=publicationFile&v=4. (Accessed: 2023-06-06)

Busch, G., Gauly, M. & Spiller, A. (2018). Opinion paper: What needs to be changed for successful future livestock farming in Europe? *Animal*, *12*(10), 1999-2001.

Daly, A., Hess, S. & Train, K. (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, *39*, 19-31.

Darby, M. R. & Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of Law and Economics*, *16*(1), 67-88.

De Bauw, M., Matthys, C., Poppe, V., Franssens, S. & Vranken, L. (2021). A combined Nutri-Score and 'Eco-Score' approach for more nutritious and more environmentally friendly food choices? Evidence from a consumer experiment in Belgium. *Food Quality and Preference*, *93*, 104276.

Dhar, R. & Simonson, I. (2003). The effect of forced choice on choice. *Journal of Marketing Research*, *40*(2), 146-160.

Diener, C., Orme, B. & Yardley, D. (2006, July). Dual response "none" approaches: Theory and practice. In *Proceedings of the sawtooth software conference* (p. 157-168).

Donato, C. & Adigüzel, F. (2022). Visual complexity of eco-labels and product evaluations in online setting: Is simple always better? *Journal of Retailing and Consumer Services*, *67*, 102961.

Godfray, H. C. J., Aveyard, P., Garnett, T., Hall, J. W., Key, T. J., Lorimer, J., ... Jebb, S. A. (2018). Meat consumption, health, and the environment. *Science*, *361*(6399).

Grömping, U. (2018). R package DoE. base for factorial experiments. *Journal of Statistical Software*, *85*, 1-41.

Grunert, K. G., Bredahl, L. & Brunsø, K. (2004). Consumer perception of meat quality and implications for product development in the meat sector–A review. *Meat Science*, *66*(2), 259-272.

Haghani, M., Bliemer, M. C., Rose, J. M., Oppewal, H. & Lancsar, E. (2021). Hypothetical bias in stated choice experiments: Part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods. *Journal of Choice Modelling*, *41*, 100322.

Heidel, A., Hagist, C. & Schlereth, C. (2021). Pricing through health apps generated data – Digital dividend as a game changer: Discrete choice experiment. *PLOS One*, *16*(7), e0254786.

Hensher, D. A. & Greene, W. H. (2003). The mixed logit model: the state of practice. *Transportation*, *30*, 133-176.

Hess, S. & Train, K. (2017). Correlation and scale in mixed logit models. *Journal of Choice Modelling*, *23*, 1-8.

Huber, J. & Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, *33*(3), 307-317.

Ikonen, I., Sotgiu, F., Aydinli, A. & Verlegh, P. W. (2020). Consumer effects of front-of-package nutrition labeling: An interdisciplinary meta-analysis. *Journal of the Academy of Marketing Science*, *48*, 360-383.

Jackson, C. (2011). Multi-state models for panel data: The msm package for R. *Journal of Statistical Software*, *38*, 1-28.

Janssen, M., Busch, C., Rödiger, M. & Hamm, U. (2016). Motives of consumers following a vegan diet and their attitudes towards animal agriculture. *Appetite*, *105*, 643-651.

Leire, C. & Thidell, Å. (2005). Product-related environmental information to guide consumer purchases – A review and analysis of research on perceptions, understanding and use among Nordic consumers. *Journal of Cleaner Production*, *13*(10-11), 1061-1070.

Majer, J. M., Henscher, H. A., Reuber, P., Fischer-Kreer, D. & Fischer, D. (2022). The effects of visual sustainability labels on consumer perception and behavior: A systematic review of the empirical literature. *Sustainable Production and Consumption*, *33*, 1-14.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*. Academic Press.

Miller, L. M. S. & Cassady, D. L. (2015). The effects of nutrition knowledge on food label use. A review of the literature. *Appetite*, *92*, 207-216.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, *78*(2), 311-329.

Newman, C. L., Burton, S., Andrews, J. C., Netemeyer, R. G. & Kees, J. (2018). Marketers' use of alternative front-of-package nutrition symbols: An examination of effects on product evaluations. *Journal of the Academy of Marketing Science*, *46*, 453-476.

Packer, J., Russell, S. J., Ridout, D., Hope, S., Conolly, A., Jessop, C., ... Croker, H. (2021). Assessing the effectiveness of front of pack labels: Findings from an online randomised-controlled experiment in a representative British sample. *Nutrients*, *13*(3), 900.

Peschel, A. O., Orquin, J. L. & Loose, S. M. (2019). Increasing consumers' attention capture and food choice through bottom-up effects. *Appetite*, *132*,

1-7.

Risius, A. & Hamm, U. (2017). The effect of information on beef husbandry systems on consumers' preferences and willingness to pay. *Meat Science*, *124*, 9-14.

Roberto, C. A., Shivaram, M., Martinez, O., Boles, C., Harris, J. L. & Brownell, K. D. (2012). The Smart Choices front-of-package nutrition label. Influence on perceptions and intake of cereal. *Appetite*, *58*(2), 651-657.

Sarrias, M. & Daziano, R. (2017). Multinomial logit models with continuous and discrete individual heterogeneity in R: The gmnl package. *Journal of Statistical Software*, *79*, 1-46.

Scarpa, R., Thiene, M. & Train, K. (2008). Utility in willingness to pay space: A tool to address confounding random scale effects in destination choice to the Alps. *American Journal of Agricultural Economics*, *90*(4), 994-1010.

Schulze, M., Spiller, A. & Risius, A. (2021). Do consumers prefer pasture-raised dual-purpose cattle when considering meat products? A hypothetical discrete choice experiment for the case of minced beef. *Meat Science*, *177*, 108494.

Sirieix, L., Delanchy, M., Remaud, H., Zepeda, L. & Gurviez, P. (2013). Consumers' perceptions of individual and combined sustainable food labels: A UK pilot investigation. *International Journal of Consumer Studies*, *37*(2), 143-151.

Song, J., Brown, M. K., Tan, M., MacGregor, G. A., Webster, J., Campbell, N. R., ... He, F. J. (2021). Impact of color-coded and warning nutrition labelling schemes: A systematic review and network meta-analysis. *PLOS medicine*, *18*(10), e1003765.

Sonntag, W. I., Lemken, D., Spiller, A. & Schulze, M. (2023). Welcome to the (label) jungle? Analyzing how consumers deal with intra-sustainability label trade-offs on food. *Food Quality and Preference*, *104*, 104746.

Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M. & de Haan, C. (2006). *Livestock's long shadow: Environmental issues and options.* Food and Agriculture Organization of the United Nations.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, *69*(3), 213-225.

Stoltenberg, B., Unfried, M. & Manewitsch, V. (2022). Better product labels for better consumer choices. *NIM Marketing Intelligence Review*, *14*(1), 49-53.

Street, D. J. & Burgess, L. (2007). *The construction of optimal stated choice experiments: Theory and methods.* John Wiley & Sons.

Talati, Z., Norman, R., Pettigrew, S., Neal, B., Kelly, B., Dixon, H., ... Shilton, T. (2017). The impact of interpretive and reductive front-of-pack labels on food choice and willingness to pay. *International Journal of Behavioral Nutrition and Physical Activity*, *14*, 1-10.

Thøgersen, J. (2000). Psychological determinants of paying attention to eco-labels in purchase decisions: Model development and multinational validation. *Journal of Consumer Policy*, *23*(3), 285-313.

Thøgersen, J., Jørgensen, A. K. & Sandager, S. (2012). Consumer decision-making regarding a "green" everyday product. *Psychology & Marketing*, *29*(4), 187-197.

Tonkin, E., Wilson, A. M., Coveney, J., Webb, T. & Meyer, S. B. (2015). Trust in and through labelling – A systematic review and critique. *British Food Journal*, *117*(1), 318-338.

Tonsor, G. T. & Wolf, C. A. (2011). On mandatory labeling of animal welfare attributes. *Food Policy*, *36*(3), 430-437.

Torma, G. & Thøgersen, J. (2021). A systematic literature review on meta sustainability labeling – What do we (not) know? *Journal of Cleaner Production*, *293*, 126194.

Train, K. E. (2009). *Discrete choice methods with simulation, Second Edition.* Cambridge University Press.

Train, K. E. & Weeks, M. (2005). Discrete choice models in preference and willingness-to-pay space. In A. Alberini & R. Scarpa (Eds.), *Applications of simulation methods in environmental and resource economics* (p. 1-16). Kluwer Academic Publishers.

Weinrich, R. & Spiller, A. (2016). Can a multi-level label do better than a binary label for animal welfare? A PLS-Analysis of Consumer Satisfaction. *International Food and Agribusiness Management Review*, *19*(1030-2016-83135), 1-30.

Xue, H., Mainville, D., You, W. & Nayga, Jr, R. M. (2010). Consumer preferences and willingness to pay for grass-fed beef: Empirical evidence from in-store experiments. *Food Quality and Preference*, *21*(7), 857-866.