Comparing preference for multi-dimensional multi-level sustainability labels of groceries across countries

Benedikt Brand
Nuremberg Institute for Market Decisions (NIM) e.V.
Birgit Stoltenberg
Nuremberg Institute for Market Decisions
Lara Dintner
University of Bayreuth

Acknowledgements:
We would like to thank the University of Bayreuth's Research Support Office for funding the panel used in these studies.

Cite as:
Brand Benedikt, Stoltenberg Birgit, Dintner Lara (2026), Comparing preference for multi-dimensional multi-level sustainability labels of groceries across countries. *Proceedings of the European Marketing Academy*, 55th, (134012)

Paper from the 55th Annual EMAC Conference, Bath, United Kingdom, June 2-5, 2026

Comparing preference for multi-dimensional multi-level sustainability labels of groceries across countries

Abstract:

Sustainability labels for groceries are one way to mitigate climate change caused by food production and consumption. However, to holistically depict "sustainability", incorporating one dimension (e.g., CO_2 emissions) as well as using binary labels (instead of encompassing multiple levels) appears insufficient to convey the granularity of sustainability with all its affected dimensions (e.g., water footprint, land use). Therefore, this study examined multi-dimensional multi-level sustainability labels (MDMLSLs) as part of a Best-Worst Scaling experiment with a representative sample of n=448 Swedish and n=454 Polish consumers to identify the one MDMLSL that best balances the trade-off between complex information and an easy-to-understand visualization. Comparing three different MDMLSL types (aggregated score only; sustainability dimension only; combined) show the highest preference for combined types of MDMLSLs. At the level of individual labels, one label proposed in academic literature (rather than one used in practice) stands out as the most/second most preferred.

Keywords: sustainability labels, food, multi-dimensional multi-level labels

Track: Social Responsibility & Ethics

1. Introduction

To counteract climate change, focusing on those issues that represent the most influential levers to do so is of particular interest. Since one quarter of anthropogenic greenhouse gases (GHG) are caused by the production of food items (Futtrup et al., 2021), transforming the food consumption patterns of eight billion people would have a tremendous impact. Moreover, the consequences of climate change, such as an increase in the intensity and number of extreme weather events (and thus, the quantity of crop failures), necessitate a more "efficient" food consumption, which can be achieved by food items with lower use of resources. This aspect becomes particularly grave given the discrepancy between an ever-growing world population on the one hand and this planet's limited resources on the other hand.

One way to successfully transform consumers' food purchase patterns is by nudging them using front-of-package labels. While multi-level front-of-package labels (such as the Nu-triScore) for indicating food items' healthiness based on sever-al dimensions are well estab-lished, no multi-dimensional multi-level sustainability label (MDMLSL) prevails in the food context yet. Currently, most labels are conceptualized binary (Weinrich & Spiller, 2016); how-ever, focusing on multi-level labels is of particular importance, since it depicts the different shades of to what extent a criterion is fulfilled more granularly (e.g., free-range breeding, dif-ferent levels of space in the barn, etc.). Most food items cannot simply be sorted into "sustain-able" and "not sustainable" and studies indicate that consumers' satisfaction for multi-level food labels is higher than for binary labels (Weinrich & Spiller, 2016).

Apart from that, incorporating more than just one dimension becomes crucial to capture the sustainability impact holistically (e.g., GHG emissions, water footprint) and thus, more re-alistically. This appears to be especially important considering the multi-facet nature of sus-tainability. However, "despite repeated calls for the development of such a holistic label, to date no label represents sustainability holistically with all of its relevant dimensions" (Futtrup et al., 2021, p. 1412). One key challenge in implementing a MDMLSL lays in translating complex sustainability information into labels that can easily by understood by every consumer (Futtrup et al., 2021).

While some studies started investigating MDMLSs, some multi-dimensional sustainability labels are merging varying sustainability aspects into one aggregate score. However, this ap-proach can dilute the potentially devastating impact of one dimension (e.g., carbon emissions) by calculating the impact of another dimension (e.g., water use) in a favourable manner. As a result, this approach provides space for environmentally harmful manipulations. In contrast, a

MDMLSL can address and solve these issues by indicating the degree of sustainability per aspect (instead of binary) as well as an aggregated score, while still outlining the impact of each dimension separately. This raises the research question evolves: How should a multi-dimensional multi-level sustainability label (MDMLSL) be conceptualized/visualized to convey many sustainability information in an easily understandable manner?

2. Theoretical Background and Literature Review

Since binary sustainability labels (e.g., Fairtrade or not) fall short of addressing all facets of sustainability for food items (e.g., GHG emissions, water use, land use, etc.), the summary of previous work is focusing on multi-level and/or multi-dimensional sustainability labels instead. In contrast to multi-dimensional sustainability labels, MDMLSLs do not only take into account various different aspects, but additional provide insights on the degree of fulfilment or lack thereof (Torma & Thøgersen, 2021). To clarify how MDMLSLs differ from multi-dimensional only or multi-level only labels, Figure 1 visually categorizes the different types of sustainability labels.

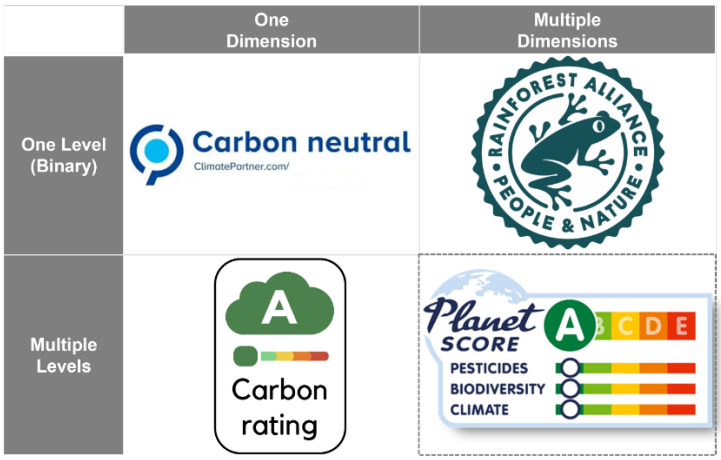


Figure 1. Categories of food labels based on levels and dimensions
(sources: Carbon rating: myemissions.co; Carbon neutral: climatepartner.com)

Focusing on multi-level sustainability labels (animal welfare), Weinrich and Spiller (2016) found that consumers show higher satisfaction for these types of labels compared to the pre-dominantly employed binary ones in Germany. Multi-level labels for carbon emissions similarly led to more sustainable choices compared to no label in Austria (Fresacher & Johnson, 2023). When focusing on sustainability labels covering multiple dimensions (without different levels), the effectiveness of these labels can be improved for some Germans, whereas these

types of labels would reduce the effect of preferring sustainably labelled products for Americans (Torma & Thøgersen, 2024). Employing qualitative expert interviews, Gröfke et al. (2021) found that the attitude towards meta sustainability labels (with various dimensions) can be altered through multi-stakeholder interactions by stakeholders from the Netherlands, Germany, and Belgium. Similarly, qualitative interviews with Danish, British, German, and Belgian experts revealed that food stakeholders would welcome a holistic sustainability label (Futtrup et al., 2021).

Comparing multi-level with multi-dimensional labels, Neumayr and Moosauer (2021) showed for German consumers that eco labels work effectively, especially those in coloured design. Examining various dimensions using multi-level labels, Hallez et al. (2021) found that sustainability labels can influence Belgium consumers to purchase more sustainability-friendly products. Similarly, combining an aggregated label score together with information on subdimensions was preferred most among Belgium consumers, which resulted in significantly higher eco-friendliness of consumers' grocery basket (Vlaeminck et al., 2014). While both multi-level and MDMLSLs work better than none, MDMLSL even perform slightly better than multi-level labels (Muller et al., 2019) among French consumers. Testing different MDMLSLs indifferent visualizations among British consumers evinced that combining a traffic light label aggregated score with subdimension indicators yields lower environmental impacts (Potter et al., 2022).

Examining the literature about multi-dimensional and/or multi-level sustainability labels in the food sector reveals four central limitations, which have not yet been addressed. First, some studies discuss different label types without testing these labels empirically (i.e., conceptual studies, such as Hélias et al., 2022; Kocsis & Kuslits, 2019; Leach et al., 2016; Lemken et al., 2021; Stein & Lima, 2022; Torma & Thøgersen, 2021). Second, some studies conduct empirical studies but fail to acquire sample sizes that allow a broader generalization of findings (e.g., Futtrup et al., 2021; Gröfke et al., 2021; Hallez et al., 2021; Muller et al., 2019; Neumayr & Moosauer, 2021; Vlaeminck et al., 2014). Third, to obtain one MDMLSL that can be used in several (European) countries, international comparisons should be conducted to increase the findings' generalizability. Fourth, some studies perform online studies only. As a result, a potential self-report bias cannot be excluded. However, given the context of sustainability, a self-report bias is very likely to occur. Therefore, this article addresses three of these four limitations of extant literature and empirically examines MDMLSLs with large-scaled samples across countries.

While MDMLSLs should cover more than one sustainability dimension, identifying the most crucial ones is a debatable endeavour. According Roa-Goyes and Pickering (2024), GHG

such as CO_2 emissions and methane are the main cause for climate change. Moreover, they are among the most important influences of groceries (Hallez et al., 2021), which is why they represent the most important share to calculate aggregated sustainability scores for food items (Vlaeminck et al., 2014). Similarly, the water footprint is among the most important dimensions (Hallez et al., 2021). Other research also considers GHG and water footprint as crucial determinants for purchasing sustainable groceries (Grebitus et al., 2016). Various studies request incorporating animal welfare (Leach et al., 2016; Stein & Lima, 2022). Furthermore, including animal welfare takes into account a more holistic approach that goes beyond an ecological perspective on sustainability. As a result, we will incorporate (i) GHG emissions, (ii) water footprint, (iii) land use, and (iv) animal welfare.

3. Method

To identify the MDMLSL for groceries that solves the trade-off between multiple information and an easy to understand conceptualization in the best possible way, all MDMLSLs currently used in industry or discussed in research are evaluated by consumers. Therefore, all MDMLSLs derived from literature (chapter 2) and industry were slightly adjusted to comprise all four dimensions – if necessary (see Table 1).

These 14 labels can be categorised into three superordinate groups. First, some MDMLSLs cover multiple dimensions, but illustrate solely an aggregated score ("Aggregated Only"). Second, other MDMLSLs outline scores for several dimensions, but omit an aggregated score ("Dimensions Only"). Third, some MDMLSLs combine both the score for each sustainability dimension together with an aggregated overall score ("Combined"). For the investigation, all labels received an equal score ("B") to make sure differences in consumers' preferences are caused by the conceptualization, not by the score itself.

A Best-Worst Scaling (BWS) experiment is conducted to find the one MDMLSL that consumers are most likely to take into account when shopping for groceries. This methodological approach proposed by Louviere and Woodworth aims at identifying the best (and worst) option among similar ones (Louviere et al., 2013). However, "best" and "worst" as the two dependent variables may be adjusted to the specific context (Brand & Kopplin, 2023). The BWS approach allows a distinct discrimination among similar items, which is especially useful when it comes to comparing the slightly different conceptualizations of sustainability labels. Using multiple choice sets with different labels and varying sets, the likelihood of choosing an option can be calculated using Hierarchical Bayes (HB) estimation.

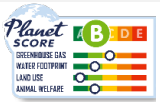
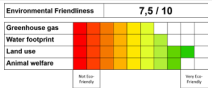
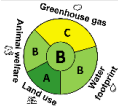
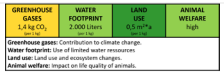
Type	Label	Source	Abbreviation
Combined		Based on Potter et al. (2022); Recommended by Stein & Lima (2022)	Flower Score
		*Industry*	Planet Score
		Based on Vlaeminck et al. (2014); Recommended by Stein & Lima (2022)	10-point Score
		Based on Stein & Lima (2022)	Pie chart
		Author's suggestion	Author's proposition
Dimensions Only		Based on Hallez et al. (2021)	Guidelines
		*Industry*; Recommended by Stein & Lima (2022)	Eaternity
		Based on Leach et al. (2016)	Table
Aggregated Only		*Industry*	Eco-Score Industry
		Based on Weinrich & Spiller (2016) and Leach et al. (2016)	Sustain-Rating
		Neumayr & Moosauer (2021)	Eco-Score Research
		*Industry*	Enviroscore
		*Industry*	Eco impact
		Potter et al. (2022)	Global

Table 1. MDMLSLs used in both studies

The BWS experiment was part of an online survey, which started with a scenario: one enters a supermarket with the intention to do groceries. Besides focusing on conventional criteria, such as price and taste, one aims at taking into account the sustainability of groceries. To do so, one will pay attention to the sustainability label to make an informed decision. To imitate decisions close to reality, the labels were not explained before the choice experiment, since consumers will usually not receive any additional label information before purchasing. The BWS experiment consisted of the 14 sustainability labels (n) derived from industry and literature. For each choice task, four labels were used (m) based on recommendations from literature. To counteract potential order biases, each label was shown three times (k). Based on the formula for calculating the number of choice sets (n/m x k), a total of eleven choice sets were generated (Brand & Kopplin, 2023).

Subsequent to the BWS, an anchor scaling (direct binary approach) was performed (Brand & Kopplin, 2023) to not only obtain relative preference (A better than B), but also absolute evaluations (label taken into account in general). Besides the BWS experiment, socio-demographic questions were asked as well as the control variables.

To counteract the limitation of low sample sizes identified in the literature review, we acquired a representative sample of n=448 Swedish and n=454 Polish consumers using an established panel provider (i.e., Kantar). We selected Swedish consumers, since they are the ones using sustainability labels the most when shopping groceries, whereas Polish consumers have the lowest levels of sustainability concerns (as part of a comparison of six European countries; Grunert et al., 2014). The surveys were translated using double-backward translation. The final sample consisted of roughly 57% females (Sweden) and 50.9% (Poland), which was caused by aiming for a 50% gender split within each generational cohort and due to the elimination of speeders. The generational cohorts match the ones in the Swedish populations with deviations below one percent, except for those 78 years or older (-1.2% deviation) and minor deviations for Polish consumers (up to 1.6% from actual society).

4. Results

The results indicate significant differences between Swedish and Polish consumers' preferences. While for Swedish consumers more information is welcomed, Poles either favour most information (Combined) or an Aggregated Only label (see Figure 2).

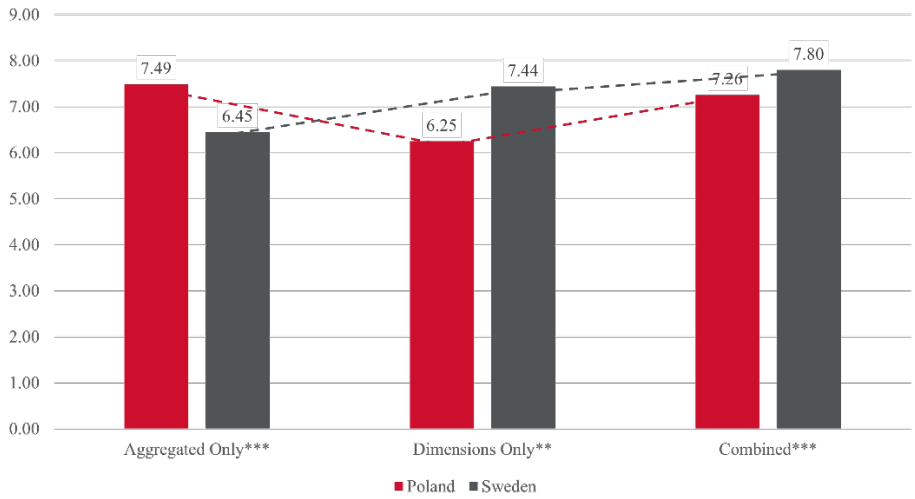


Figure 2. Relative preference for Aggregated Only, Dimensions Only and Combined for Polish versus Swedish consumers (Note: **p<0.01; ***p<0.001)

The individual label results show that the MDMLSL proposed by Potter et al. (2022) is preferred most ("Flower-Score") among Swedes and second-most among Poles (Figure 3). In general, the group of MDMLSLs that combine both an aggregated score and displaying dimensions tend to perform better than those with an aggregated score only or those without any agg-

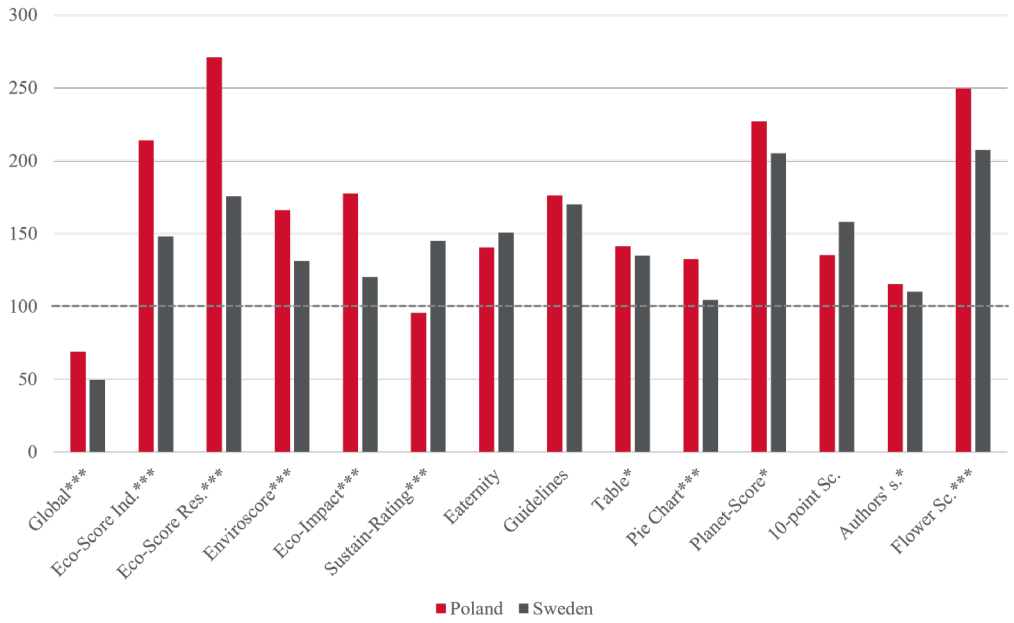


Figure 3. Absolute preference for MDMLSLs among Poles and Swedes with anchor of '100' (Note: *p<0.05; **p<0.01; ***p<0.001)

regated score. Accordingly, out of the five most favoured MDMLSLs, three represent such combined solutions (i.e., Flower-Score, Planet-Score, 10-point Score) among Swedes. Apart

from that, almost all labels are taken into account by Swedish and Polish consumers when shopping for groceries. Only the global label tested by Potter et al. (2022) will not be considered ("Aggregated Only").

5. Discussion

To identify the one MDMLSL that handles the trade-off between many, complex information on the one hand and being easily understandable on the other hand, this article exploratively investigates consumers' preferences of several MDMLSLs for food items. Since sustainability is not binary and constitutes of several different aspects (such as GHG emissions, water footprint), this study's focus lays on investigating MDMLSLs. Accordingly, those MDMLSL are favoured, which combine displaying both an aggregated score and additional dimensions ("Combined"). In particular, one label proposed by research (i.e., Flower-Score; Potter et al., 2022) yields the highest preference.

As a result, this study corroborates the positive effect of the Flower-Score on consumers' choices (Potter et al., 2022), as well as the positive evaluation by Stein and Lima (2022). Moreover, it confirms that sustainability labels with multiple dimensions and higher information density perform better (Muller et al., 2019). However, this contradicts findings from previous literature claiming that consumers prefer simple sustainability labels (Roa-Goyes & Pickering, 2024). This discrepancy may stem from cultural differences (investigating Canadian and Argentinian consumers) compared to European countries (here: Sweden and Poland), respondents' age (18-24 years versus 18-87 years), or a combination of both.

This study responds to the call for defining which sustainability dimensions should be incorporated in MDMLSLs for groceries (Stein & Lima, 2022). To reveal the one label consumers are most likely to take into account when shopping groceries, this article contributes by summarizing the labels used in practice and those discussed in literature to disentangle the current "label jungle". As a result, one specific label (Flower Score) stands out as being the most preferred. Additionally, the "Combined" group of MDMLSLs is consumers' most favourite ones in Sweden and as also welcomed among Poles with the Aggregated Only labels slightly ahead.

Apart from that, this article responds to the identified lack of empirical research on the effectiveness of environmental labels with varying levels of complexity (Hallez et al., 2021). Moreover, the findings revealed answers the call for the development of a holistic sustainability label with all relevant dimensions (Futtrup et al., 2021).

References

Brand, B. M., & Kopplin, C. S. (2023). Effective Return Prevention Measures in the Post-purchase Stage: A Best-Worst Scaling Approach. *Marketing ZFP*, *45*(1), 30–47.

Fresacher, M., & Johnson, M. K. P. (2023). Designing climate labels for green food choices. *Journal of Cleaner Production*, *430*, 139490.

Futtrup, R., Tsalis, G., Pedersen, S., Dean, M., Benson, T., & Aschemann-Witzel, J. (2021). Is the whole more than the sum of its parts? Challenges and opportunities for a holistic consumer-friendly sustainability label on food. *Sustainable Production and Consumption*, *28*, 1411–1421.

Grebitus, C., Steiner, B., & Veeman, M. M. (2016). Paying for sustainability: A cross-cultural analysis of consumers' valuations of food and non-food products labeled for carbon and water footprints. *Journal of Behavioral and Experimental Economics*, *63*, 50–58.

Gröfke, N., Duplat, V., Wickert, C., & Tjemkes, B. (2021). A multi-stakeholder perspective on food labelling for environmental sustainability: Attitudes, perceived barriers, and solution approaches towards the "traffic light index". *Sustainability*, *13*(2), 933.

Grunert, K. G., Hieke, S., & Wills, J. (2014). Sustainability labels on food products: Consumer motivation, understanding and use. *Food Policy*, *44*, 177–189.

Hallez, L., Qutteina, Y., Boen, F., & Smits, T. (2021). The abc's of ecological and nutrition labels. The impact of label theme and complexity on the environmental footprint of online grocery choices. *Sustainability*, *13*(5), 2474.

Hélias, A., van der Werf, H. M. G., Soler, L.-G., Aggeri, F., Dourmad, J.-Y., Julia, C., Nabec, L., Pellerin, S., Ruffieux, B., & Trystram, G. (2022). Implementing environmental labelling of food products in France. *The International Journal of Life Cycle Assessment*, *27*(7), 926–931.

Kocsis, T., & Kuslits, B. (2019). Multidimensional labelling: Closing the sustainability information gap between producers, consumers and sustainability science in the food sector. *Periodica Polytechnica Social and Management Sciences*, *27*(1), 9–16.

Leach, A. M., Emery, K. A., Gephart, J., Davis, K. F., Erisman, J. W., Leip, A., Pace, M. L., D'Odorico, P., Carr, J., & Noll, L. C. (2016). Environmental impact food labels combining carbon, nitrogen, and water footprints. *Food Policy*, *61*, 213–223.

Lemken, D., Zühlsdorf, A., & Spiller, A. (2021). Improving consumers' understanding and use of carbon footprint labels on food: proposal for a climate score label. *EuroChoices*, *20*(2), 23–29.

Louviere, J., Lings, I., Islam, T., Gudergan, S., & Flynn, T. (2013). An introduction to the application of (case 1) best–worst scaling in marketing research. *International Journal of Research in Marketing*, *30*(3), 292–303.

Muller, L., Lacroix, A., & Ruffieux, B. (2019). Environmental labelling and consumption changes: a food choice experiment. *Environmental and Resource Economics*, *73*, 871–897.

Neumayr, L., & Moosauer, C. (2021). How to induce sales of sustainable and organic food: The case of a traffic light eco-label in online grocery shopping. *Journal of Cleaner Production*, *328*, 129584.

Potter, C., Pechey, R., Clark, M., Frie, K., Bateman, P. A., Cook, B., Stewart, C., Piernas, C., Lynch, J., & Rayner, M. (2022). Effects of environmental impact labels on the sustainability of food purchases: Two randomised controlled trials in an experimental online supermarket. *PLoS One*, *17*(11), e0272800.

Roa-Goyes, S., & Pickering, G. J. (2024). Promoting a sustainable diet through carbon labeling of food: Insights from young consumers in the Americas. *Sustainable Production and Consumption*, *44*, 179–187.

Stein, A. J., & Lima, M. de (2022). Sustainable food labelling: considerations for policy-makers. *Review of Agricultural, Food and Environmental Studies*, *103*(2), 143–160.

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

Torma, G., & Thøgersen, J. (2024). Can a meta sustainability label facilitate more sustainable consumer choices? *Business Strategy and the Environment*, *33*(2), 283–306.

Vlaeminck, P., Jiang, T., & Vranken, L. (2014). Food labeling and eco-friendly consumption: Experimental evidence from a Belgian supermarket. *Ecological Economics*, *108*, 180–190.

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.