

Absolutely unhealthy, relatively healthy - The effect of consumer reference points and loss aversion in evaluating front-of-package nutrition labels

David Olk

University of Groningen

Koert van Ittersum

University of Groningen

Tammo Bijmolt

University of Groningen

Cite as:

Olk David, van Ittersum Koert, Bijmolt Tammo (2023), Absolutely unhealthy, relatively healthy - The effect of consumer reference points and loss aversion in evaluating front-of-package nutrition labels. *Proceedings of the European Marketing Academy*, 52nd, (114090)

Paper from the 52nd Annual EMAC Conference, Odense/Denmark, May 23-26, 2023



Absolutely unhealthy, relatively healthy - The effect of consumer reference points and loss aversion in evaluating front-of-package nutrition labels

Abstract

Previous research has found that front-of-package labels often fail to increase healthy purchases. More importantly, unhealthy purchases are often not reduced and, in some cases, even increased. In this paper, we introduce the concept of reference points to the labelling domain as an underlying mechanism. Specifically, we argue that products can be perceived as relatively healthy compared to expectations (Internal Reference Point) or other products (Contextual Reference Point), even when they are not healthy in absolute terms. In three experiments, we analyze the nature of those reference points, show that loss aversion applies, and find evidence that consumers tend to choose products that healthier than the reference points.

Keywords: Healthy food purchases, front-of-package labels, reference points

Conference track: Consumer Behavior

1. Introduction

Health issues related to food consumption pose an increasing threat to people worldwide (WHO, 2018). In response, so-called front-of-package labels (FOPLs) were introduced to provide a prominent and easily understandable source of nutritional information (Chantal & Hercberg, 2017). FOPLs visualize or summarize the nutritional content of a product using simpler graphics, colors, and/or summary scores (Ikonen et al., 2020). Researchers have found mixed results regarding the effectiveness of FOPLs on the healthiness of purchases (Crocker et al., 2020; Ikonen et al., 2020). Additionally, while the findings show that FOPLs can increase the purchasing of healthy products, the health perception and purchasing of unhealthy products is often not reduced and even increases in some studies (Ikonen et al., 2020).

In this paper, we argue that the increase of unhealthy purchases after the introduction of a FOPL can be explained using reference-dependent theory (Tversky & Kahneman, 1991). We argue that consumers form an estimation for the healthiness of each product category (i.e., a reference point). A reference point is an anchor for the healthiness of a product that serves as a baseline for comparison when making a product choice (Briesch et al., 1997; Tversky & Kahneman, 1991). We argue that consumers do not evaluate the information provided by the introduction of a FOPL in absolute terms, but relative to their reference points. These relative evaluations make consumers perceive health gains or losses, that can make a product that is unhealthy in absolute terms appear to be relatively healthy and vice-versa. The aim of this research is thus to gain a better understanding of consumers' health reference points, and to analyze how these reference points can be used to explain and predict choice behavior.

The main scientific contributions of this paper are that we (1) offer a missing understanding about the asymmetric effects of health-based FOPL on healthy versus unhealthy products, (2) integrate reference-dependent theory into the health intervention domain, and (3) provide novel and unique insights into the existence of internal and Contextual Reference Points with regards to the healthiness of products.

2. Theoretical framework

2.1 Front-of-package Labels

There is an ongoing discussion regarding the effectiveness of FOPLs in motivating healthier purchases. Researchers generally agree that FOPLs are effective in increasing consumers' ability to identify healthier products (Ikonen et al., 2020). However, previous meta-analyses find that effects on purchase behavior are often not significant (Crocker et al., 2020) or small (Ikonen et al., 2020). Additionally, while health interventions generally often fail to

reduce the purchasing unhealthy products, FOPLs even increase unhealthy purchases in some cases (Ikonen et al., 2020). To the best of our knowledge, this unusual but detrimental side effect has not been explained in previous literature.

While the main focus of this paper is to provide an explanation for these unexpected side effects of health labels, we also contribute to the ongoing general discussion of whether health labels work in general. Given that there is predominantly support for the positive effect of FOPLs on the healthiness of the shopping basket, we take an optimistic approach and hypothesize in favor of the FOPL's effectiveness while keeping the discussion above in mind.

H1a: Consumers who are exposed to a front-of-package health label purchase healthier products on average than consumers who are not exposed to the label.

H1b: For consumers who are exposed to the label, increases in the front-of-package label's healthiness rating will lead to higher choice probabilities.

2.1 Internal and Contextual Reference Points

Reference points are mental anchors that are formed based on previous exposures to explicit information or subjective experiences with a product (Kalyanaram & Winer, 1995). Reference-dependent theory (Tversky & Kahneman, 1991) states that when people are exposed to new information or faced with a new choice option, they compare it to their reference point. Past research on labeling has mainly considered the direct relationship between FOPLs and healthiness of purchases. However, De Temmerman et al. (2021) show that this effect is mediated by consumers' perceived healthiness of a product. Therefore, we argue that it is important to consider consumers' reference points and how they influence perceived product healthiness to explain their choice behavior. Based on past research, we consider the Internal Reference Point as suggested in reference-dependent theory (Tversky & Kahneman, 1991), and the Contextual Reference Point as suggested in price literature (Briesch et al., 1997).

The Internal Reference Point is a subjective, memory-based anchor (Tversky & Kahneman, 1991). Consumers initially form and then update their Internal Reference Points with each exposure to a product or information about the product. Previous research finds that consumers often neglect the currently existing nutritional information on the back of the package (Graham et al., 2012), and are not able to estimate nutrients accurately (Burton et al., 2009). Therefore, we expect that consumers Internal Reference Points regarding the healthiness of a product category will differ from the actual product healthiness.

In addition to their Internal Reference Point, consumers can form supplementary reference points based on further external cues (Tversky & Kahneman, 1991). Therefore, we

also include the Contextual Reference Point as an external, stimulus-based reference point (Briesch et al., 1997). It is argued that Contextual Reference Points are (temporarily) formed when consumers do not clearly recall their Internal Reference Point at the point of purchase, as it is not formed yet or not sufficiently salient in that moment (Briesch et al., 1997). In these situations, consumers use the input from other products (i.e., their actual healthiness based on FOPLs) to form a reference point for the healthiness of a product.

As the Internal Reference Point is based on consumers' subjective beliefs and the Contextual Reference Point is based on factual information, they can differ from each other. In this case, previous literature suggests the Internal Reference Point is dominant as consumers hold on to longer-held beliefs due to the Status Quo Bias (Tversky & Kahneman, 1991). Furthermore, it is argued that Contextual Reference Points are only formed when an Internal Reference Point is non-existent or not very salient (Briesch et al., 1997). Based on this discussion, we formulate two hypotheses:

H2: Consumers' Internal Reference Points significantly differ from the Contextual Reference Points.

H3: The Internal Reference Point has a better predictive ability for consumer choices than the Contextual Reference Point.

2.3 Health gains and losses

In the process of comparing the health information of a product to the reference point, consumers can experience increases in value (i.e., a health gain) or decreases in value (i.e., a health loss) (Burton et al., 2009; Tversky & Kahneman, 1991). For the Internal Reference Point, this means that an actual product's healthiness can be perceived as better or worse compared to the consumers' individual, subjective anchor. For the Contextual Reference Point, consumers can find that a product is performing better or worse compared to the product alternatives available. If consumers experience a health gain, they should be more likely to choose a certain product and vice-versa with a health loss. Therefore, we hypothesize for both the Internal Reference Point (H4a and H4b) and the Contextual Reference Point (H5a and H5b) that a health gain leads to an increase in choice likelihood and a health loss leads to a decrease in choice likelihood.

H4a: A health gain based on the consumer's Internal Reference Point leads to an increase in choice likelihood.

H4b: A health loss based on the consumer's Internal Reference Point leads to a decrease in choice likelihood.

H5a: A health gain based on the consumer's Contextual Reference Point leads to an increase in choice likelihood.

H5b: A health loss based on the consumer's Contextual Reference Point leads to an increase in choice likelihood.

2.4 Loss aversion

Regarding the magnitude of the effects of gains and losses on choice likelihood, reference dependent theory also posits that consumers experience loss aversion (Tversky & Kahneman, 1991). This means that people assign substantially more weight to a (potential) health loss than to an equal (potential) health gain. As a result, the decrease in purchase likelihood of a product that is labeled as less healthy than expected is higher than the increase in purchase likelihood for a product that is healthier than expected. We hypothesize this to be the case for both the Internal Reference Point (H6a) and the Contextual Reference Point (H6b).

H5a: For the Internal Reference Points, the negative effect of a health loss (H3b) is larger than the positive effect of a health gain (H3a).

H5b: For the Contextual Reference Points, the negative effect of a health loss (H5b) is larger than the positive effect of a health gain (H5a).

3. Empirical study design

We test our hypotheses in three experimental studies. For all studies, we chose for the Nutri-Score as the operationalization of a summary score-type FOPL and, thereby, as the measure for the healthiness of a product. The Nutri-Score scores each product on a colored scale from A (very good nutritional value, dark-green) to E (very bad nutritional value, dark-red) based on a formula that combines the amounts of key nutrients and calories in a product (Chantal & Hercberg, 2017). The label has been recommended by previous researchers (De Temmerman et al., 2021).

3.1 Study 1 - Consumers' Internal and Contextual Reference Points

In Study 1, we examine whether the participants' Internal Reference Points significantly differ from the location of the theoretical Contextual Reference Point per category, as represented by the average Nutri-Score. For that purpose, participants are asked to report their Internal Reference Points by assigning a Nutri-Score to eight product categories. Participants are given ten exemplary products per category. When participants have to choose a product, the Nutri-Scores of all those products would serve as the basis for forming the Contextual

Reference Point. Therefore, we set the Contextual Reference Point equal to the average Nutri-Score of those products. For this study, we recruited a small sample of Dutch participants (n=53) via Prolific to participate in a Qualtrics study (31 males, $M_{age} = 28,58$, $SD_{age} = 7.37$). We test whether the participants' Internal Reference Points significantly differ from the Contextual Reference Point using a t-test per product category (see Table 1).

Table 1

Distribution of participants Internal Reference Points and two-tailed t-test results for comparing the Internal Reference Point with the Contextual Reference Point (average category Nutri-Score)

Internal Reference Point	Chips	Ice-cream	Fruit	Snacks	Yogurt	Sodas	Baked Goods	Soups
A (5)	0	0	27	0	2	1	2	0
B (4)	0	1	24	1	33	2	20	15
C (3)	7	10	1	6	16	7	22	27
D (2)	25	25	0	22	1	27	6	10
E (1)	21	17	1	24	1	16	3	1
Average Internal Reference Point	1.74	1.91	4.43	1.70	3.64	1.96	3.23	3.06
Contextual Reference Point	2.90	2.40	4.80	1.50	4.10	2.90	3.70	3.70
Difference	-1.16	-0.49	-0.37	0.20	-0.46	-0.94	-0.47	-0.64
t	-12.152	-4.507	-3.989	2.034	-5.081	-7.582	-3.588	-6.280
df	51	51	51	51	51	51	51	51
p-value	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***
LCL	1.560	1.711	4.349	1.503	3.531	1.737	3.028	2.903
UCL	1.940	2.136	4.651	1.920	3.853	2.224	3.510	3.289

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Across all product categories, we find that the sample of participants' Internal Reference Points significantly differs from the Contextual Reference Points. This is strong support for Hypothesis 2. Except for the product category snacks, consumers estimate products (Internal Reference Point) as unhealthier than they actually are (Contextual Reference Points).

3.2 Study 2 – Loss aversion in reacting to health gains and losses

To test how people react to unexpected health gains or losses in a product, we recruited a Dutch convenience sample of 141 people to participate in a Qualtrics study (55 males, $M_{age} = 36.24$, $SD_{age} = 16.70$). This experiment was conducted as a 2 (healthy versus unhealthy product) x 2 (healthy versus unhealthy FOPL score) between-subjects design. Specifically, we created product pairs from three product categories (bread, milk, and salty snacks) with one rather healthy and one rather unhealthy option, and labeled them with the rather healthy Nutri-Score B or rather unhealthy Nutri-Score D. As a result, participants saw products that were either as expected, healthier than expected, or unhealthier than expected. The dependent variable is the consumers' willingness to buy each product as an indication for the value they assign to it.

Our expectation for this experiment is that, due to loss aversion, the decrease in willingness to buy in the case of the health loss condition is larger than the increase in willingness to buy in the health gain condition. Our results indicate that the interaction between the product healthiness and labeled healthiness is significant ($F = 4.886$, $p = 0.029$). Specifically, we see that while there is a

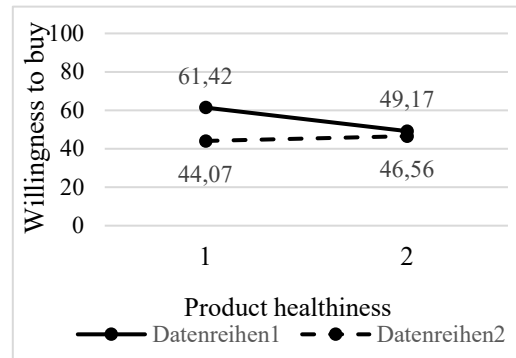


Figure XXX: WTB per condition

substantial decrease in willingness to buy for the healthy product that comes with a health loss, there is not a big difference between the unhealthy product regardless of the labelling. This indicates that consumers react more heavily to a health loss than to a health gain and provides a first proof of concept for Hypotheses 5a and 5b.

3.1 Study 3

For Study 3, we recruited 675 participants from Prolific to fill in a Qualtrics questionnaire (283 males, $M_{age} = 28.09$, $SD_{age} = 0.35$). Participants were initially asked to report their Internal Reference Points as in Study 1. Thereafter, they complete a virtual shopping trip by making a product choice for nine product categories. For each category, each participant receives a random draw of six products out of a pool of twelve potential products and a no-choice option. By doing so, a range of Contextual Reference Points were generated, as varying choice sets result in different average Nutri-Scores.

The analysis is split into two parts. The first part (Study 3A) considers the entire dataset and analyzes the effect of introducing a FOPL during the purchasing of products. Our expectations based on previous research are that the FOPLs generally have a small positive effect, although there are some mixed findings. The results of study 3A indicate that there is a significant difference ($F = 34.681$, $p < 0.001$) between people that had the FOPL present ($\bar{x}_{FOPL} = 3.44$, with NS A = 5 and NS E = 1) compared to the ones who did not have it present ($\bar{x}_{Control} = 3.25$). This supports Hypothesis 1a.

The second part (Study 3B) focuses on the half of the participants who are in the experimental condition where the FOPL was shown. With this subsample, we analyze the effect of reference points on the choice behavior of the participants. Specifically, we estimate three models with the same data and same control variables but with differing main independent variables, namely: 1) the Nutri-Score, 2) the Nutri-Score – Internal Reference Point, and 3) the

Nutri-Score – Contextual Reference Point. For each model, we test how the various levels of the independent variables (in nominal form) influence the choice likelihood of products.

For the *Nutri-Score model*, we expect that a higher Nutri-Score leads to a higher choice likelihood. Despite a few exceptions, the results (see Table 2) generally approve this expectation as healthier Nutri-Scores have a positive effect on choice likelihood and less healthy Nutri-Scores have a negative effect on choice likelihoods. This provides additional support for Hypothesis 1b.

Table 2

Results of the Nutri-Score minus Internal Reference Point model with nominal effects

Effect	Chips	Ice-cream	Sweet snacks	Yogurt	Sodas	Baked goods	Soups	Cereal	Pizza
Nutri-Score									
NS A	n.a.	n.a.	n.a.	0.559	2.26	0.832	0.528	1.093	n.a.
NS B	1.028	-0.006	-0.4435	-0.479	0.155	0.618	0.180	0.3213	0.287
NS C	0.248	0.530	n.a.	-0.080	0.067	-1.049	-0.708	-0.404	-0.443
NS D	-1.276	-0.013	1.163	n.a.	-1.650	-0.400	n.a.	-1.010	0.155
NS E	n.a.	-0.511	-0.720	n.a.	-0.832	n.a.	n.a.	n.a.	n.a.
p-value	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

For the *Nutri-Score – Internal Reference Point model*, we expect that health gains (products being healthier than expected) should increase the purchase likelihood, while health losses (products being less healthy than expected) should decrease the purchase likelihood. The results indicate an almost perfectly consistent pattern in line with our expectations, providing support for Hypothesis 4a and 4b (see Table 3). This indicates that independent of a product’s actual healthiness, customers are more likely to choose it as long as it is healthier than their expectations.

Table 3

Results of the Nutri-Score minus Internal Reference Point model with nominal effects

Effect	Chips	Ice-cream	Sweet snacks	Yogurt	Sodas	Baked goods	Soups	Cereal	Pizza
NS-Internal Reference Point									
Big gain	3,049	0,568	-0,193	0,177	1,256	0,815	1,051	1,317	0,550
Small gain	1,783	0,507	1,296	0,679	0,529	0,725	0,761	0,574	0,076
As expected	0,806	-0,055	-0,031	0,301	-0,275	-0,260	0,017	0,056	-0,224
Small loss	0,064	-0,035	-0,311	-0,109	-0,312	-0,421	-0,468	-0,911	-0,397
Big loss	-5,702	-0,985	-0,761	-1,049	-1,198	-0,859	-1,361	-1,035	-0,005
p-value	<0.001 ***	0.006 **	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	0.037 *

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Finally, for the *Nutri-Score - Contextual Reference Point model*, we also expect that products at a health gain (product being healthier than the average of the offered products)

should increase choice likelihood and vice-versa for health losses. We generally observe a similar trend here as with the previous model (see Table 4). This indicates that consumers are overall more likely to choose products when they are healthier than the average and less likely to purchase products that are less healthy than the average product. It should be noted that four out of the 42 effects are exceptions to the general pattern.

Table 4

Results of the Nutri-Score minus Contextual Reference Point model with nominal effects

Effect	Chips	Ice-cream	Sweet snacks	Yogurt	Sodas	Baked goods	Soups	Cereal	Pizza
NS-Contextual Reference Point									
Big gain	1,060	0,052	-0,279	1,183	1,394	1,021	0,737	0,874	n.a. †
Small gain	0,210	0,634	1,329	0,187	0,253	0,733	0,188	0,546	0,306
As expected	0,034	-0,134	-0,489	-1,037	-0,146	-0,560	0,470	0,070	-0,202
Small loss	-1,303	0,046	-0,561	-0,118	-0,805	-1,155	-0,585	-0,534	-0,443
Big loss	n.a. †	-0,598	n.a. †	-0,216	-0,696	-0,040	-0,809	-0,957	0,339
p-value	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

† Event did not occur within the dataset.

When comparing, the model fits of the three different models for all nine categories, we find that the *Nutri-Score – Contextual Reference Point* model performs best for most of the categories, closely followed by the *Nutri-Score model*. The *Nutri-Score – Internal Reference Point* model is outperformed in every product category. This is contrary to Hypothesis H3.

4. Discussion

In this paper, we suggest an underlying mechanism that can explain the findings in previous literature that the introduction of health FOPL does not reduce and, in some cases, even increases the purchasing of unhealthy products (Ikonen et al., 2020). In our three studies, we generate extensive insights regarding the nature of consumers' Internal Reference Points, and show that they differ from both the Contextual Reference Point and the products' actual healthiness. Most importantly, our analyses of Study 3B find that consumers are more likely to choose products that are labeled as healthier than the consumers' Internal Reference Point or the Contextual Reference Points, while they are less likely to choose products that are labeled as less healthy than the reference points. This can explain that once labels are introduced, consumers can perceive products as relatively healthy, even though they are unhealthy in absolute terms. Further discussion of the results, implications, and limitations will be presented at the conference.

5. References

- Briesch, R. A., Krishnamurthi, L., Mazumdar, T., & Raj, S. P. (1997). A Comparative Analysis of Reference Price Models. *Journal of Consumer Research*, 24(2), 202–214. <https://doi.org/10.1086/209505>
- Burton, S., Howlett, E., & Tangari, A. H. (2009). Food for Thought: How Will the Nutrition Labeling of Quick Service Restaurant Menu Items Influence Consumers' Product Evaluations, Purchase Intentions, and Choices? *Journal of Retailing*, 85(3), 258–273. <https://doi.org/10.1016/j.jretai.2009.04.007>
- Chantal, J., & Hercberg, S. (2017). Development of a new front-of-pack nutrition label in France: the five-colour Nutri-Score. *Public Health Panorama*, 03(04), 712–725.
- Crocker, H., Packer, J., Russell, S. J., Stansfield, C., & Viner, R. M. (2020). Front of pack nutritional labelling schemes: a systematic review and meta-analysis of recent evidence relating to objectively measured consumption and purchasing. *Journal of Human Nutrition and Dietetics*, 33(4), 518–537. <https://doi.org/10.1111/jhn.12758>
- De Temmerman, J., Heeremans, E., Slabbinck, H., & Vermeir, I. (2021). The impact of the Nutri-Score nutrition label on perceived healthiness and purchase intentions. *Appetite*, 157(February 2021), 104995. <https://doi.org/10.1016/j.appet.2020.104995>
- Graham, D. J., Orquin, J. L., & Visschers, V. H. M. (2012). Eye tracking and nutrition label use: A review of the literature and recommendations for label enhancement. *Food Policy*, 37(4), 378–382. <https://doi.org/10.1016/j.foodpol.2012.03.004>
- Ikonen, I., Sotgiu, F., Aydinli, A., & Verlegh, P. W. J. (2020). Consumer effects of front-of-package nutrition labeling: an interdisciplinary meta-analysis. *Journal of the Academy of Marketing Science*, 48(3), 360–383. <https://doi.org/10.1007/s11747-019-00663-9>
- Kalyanaram, G., & Winer, R. S. (1995). Empirical Generalizations from Reference Price Research. *Marketing Science*, 14(3_supplement), G161–G169. <https://doi.org/10.1287/mksc.14.3.g161>
- Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, 106(4), 1039–1061. <https://doi.org/10.2307/2937956>
- WHO. (2018). *Noncommunicable diseases*. <https://www.who.int/news-room/factsheets/detail/noncommunicable-diseases>