Visual Attention to Front-of-Pack Nutrition Labels:

An Eye-Tracking Comparison of Simplified and Detailed Health Cues

Masami Asakawa[a]* and Masao Okano[b]

[a]*Faculty of Health and Nutrition, Bunkyo University, Kanagawa, Japan;*

[b]*Faculty of Information and Systems, Bunkyo University, Kanagawa, Japan*

Masami Asakawa is a Professor at the Faculty of Health and Nutrition, Bunkyo University, Japan. Her research focuses on consumer behavior, health communication, and nutrition labeling. Masao Okano is a Professor at the Faculty of Information and Systems, Bunkyo University, Japan. His research interests include consumer behaviour, health communication, and data analysis.

Correspondence: Masao Okano, Faculty of Information and Communication, Bunkyo University, Kanagawa, Japan. Email: okano@shonan.bunkyo.ac.jp; Masami Asakawa, Faculty of Health and Nutrition, Bunkyo University, Kanagawa, Japan. Email: asamasa@shonan.bunkyo.ac.jp

Visual Attention to Front-of-Pack Nutrition Labels:

An Eye-Tracking Comparison of Simplified and Detailed Health Cues

Abstract

This study investigates how two types of Front-of-pack nutrition labels (FOPNL) — Nutri-Score, a summary-evaluation label, and Multiple Traffic Lights (MTL), a nutrient-specific label—differ in eliciting visual attention among Japanese consumers, for whom empirical evidence remains scarce given the absence of an FOPNL system in Japan. A between-subjects eye-tracking experiment with university students yielded a final analytic sample of 29 participants (Nutri-Score: n = 12; MTL: n = 17). Participants viewed five dairy-category package stimuli spanning healthiness levels A through E, and time to first fixation and total fixation count on the label area were analyzed using Mann–Whitney U tests with Bonferroni correction ($\alpha = .01$).

Nutri-Score was detected significantly faster than MTL only for the healthiest product (level A). In contrast, Nutri-Score elicited significantly fewer total fixations than MTL in four of the five conditions, with medium-to-large effect sizes ($r = .48–.69$), indicating a robust reduction in processing load. These findings suggest that judgments based on Nutri-Score tend to be completed through heuristic processing alone, whereas MTL appears to invite not only heuristic processing but also more elaborative processing involving the reference to and integration of individual nutrient cues.

Keywords: Front-of-pack nutrition labels; Nutri-Score; Multiple Traffic Lights; eye-tracking; visual attention; health communication

Introduction

Against the backdrop of dietary health challenges in Japan, there is growing demand for simple and accessible nutritional information. Because conventional nutrition fact panels require high levels of literacy and are of limited use in actual purchasing environments (Roberto & Khandpur, 2014; van Herpen & van Trijp, 2011), the importance of intuitive Front-of-pack nutrition labels (FOPNL) has been increasing.

Everyday food purchasing is typically characterized by strong time constraints and a large number of options. Under such conditions, consumers often rely on heuristic processing (a decision-making approach that prioritizes efficiency over accuracy) rather than carefully reading and comparing nutritional information. FOPNL systems are designed to convey nutritional quality in a simplified and intuitive manner, enabling consumers to assess the healthiness of foods at a glance even under conditions of limited time and attention.

FOPNL systems can be broadly classified into two types. The first category consists of summary-evaluation labels, such as Nutri-Score and the Healthier Choice Logo, which provide a single integrated assessment of a product's nutritional quality. The second category consists of nutrient-specific labels, such as the Multiple Traffic Lights (MTL) system, which presents information for individual nutrients using evaluative cues (Andrews et al., 2011; Egnell et al., 2018). Although both share the feature of being "simplified labels placed on the front of the package," they differ qualitatively in their degree of information

aggregation. The present study asks how this difference manifests in consumers' visual attention and judgment.

Partly because no Japanese FOPNL system yet exists, empirical research on how the two types of FOPNL elicit different visual attention among Japanese consumers remains limited. This study aims to fill this gap using eye-tracking methodology.

Eye-Tracking Studies on Nutrition Labels and Hypotheses

The Elaboration Likelihood Model (ELM) of Petty & Cacioppo (1986) distinguishes responses to persuasive communication into a central route and a peripheral route, depending on the recipient's motivation and ability. In the central route, the recipient scrutinizes the message content and judges it logically. In the peripheral route, the recipient judges intuitively based on peripheral cues of the message (such as color, symbols, and logos).

Conventional tabular nutrition labels presuppose central-route processing involving the reading and comparison of numerical values. However, in actual purchasing environments with strong time constraints, consumers tend to rely on intuitive cues rather than detailed information, so peripheral-route processing becomes dominant. FOPNL can be considered to be designed as effective stimuli for this peripheral route. That said, FOPNL does not elicit a single mode of processing; the depth of processing required can differ depending on label format. Summary-evaluation labels (Nutri-Score), by aggregating multiple pieces of nutritional information into a single indicator, may promote processing closer to the peripheral route. By contrast, nutrient-specific labels (MTL), which present color-coded information for each nutrient, require item-by-item reference and

integration, and may demand relatively more elaborative processing. This study examines how this difference in processing depth is reflected in indicators of visual attention.

Prior Research and Derivation of Hypotheses

Existing eye-tracking research on nutrition labels has been conducted mainly in Europe and North America, and conclusions about which type of FOPNL attracts more attention have not necessarily converged (see Table 1).

Insert Table 1 here

Table 1. Prior Research on the Comparison of Nutrition Labels Based on Eye-Tracking Experiments

Inconsistencies among prior studies are largely attributable to differences in target foods and task settings; when the focus is narrowed to the visual detection properties of the label format itself, reports supporting the detection advantage of logo-type labels constitute the central finding. van Herpen & van Trijp (2011) report that logo-type labels corresponding to summary-evaluation labels such as Nutri-Score are detected more quickly than nutrient-specific labels such as MTL. Based on this finding, we set out the following hypotheses. Because the present study also manipulates the healthiness level of foods, differences across healthiness levels are likewise examined.

Hypothesis 1: Time to first fixation is shorter for Nutri-Score than for MTL.

Note that eye-tracking indices capture two qualitatively distinct aspects. First, time to first fixation reflects how quickly a label is visually detected—that

is, its attentional capture. Second, total fixation count reflects the amount of processing required to extract information from the label. These are separate constructs and are in principle compatible: a given label may be "detected quickly and processed to completion in a short time." The property of summary-evaluation labels—aggregating information into a single symbol—is precisely what predicts this compatibility.

In addition, van Herpen & van Trijp (2011) show that logo-type labels generally require less visual attention than nutrient-specific labels. Gabor, Stojnić, & Ostić (2020) likewise report that Nutri-Score requires fewer visual resources than MTL. Based on these findings, we propose the following hypothesis.

Hypothesis 2: Nutri-Score elicits a smaller total fixation count than MTL.

Methods

From November to December 2022, in a dedicated room within the university, the experiment was conducted individually with 36 university students, one participant at a time. A between-subjects design was adopted: 18 participants viewed five product packages with Nutri-Score labels, and 18 viewed packages with MTL labels. Due to severe myopia or reflective eye makeup, seven participants failed eye-tracking calibration or had gaze data of insufficient accuracy and were excluded. The breakdown was six in the Nutri-Score group and

one in the MTL group, yielding a final analytic sample of N = 29 (Nutri-Score group: 18→12; MTL group: 18→17).

The unequal group sizes and the overall limited sample size constitute important constraints on the interpretation of the results (discussed later). Demographic characteristics of the participants are summarized in Table 2.

Insert Table 2 here

Table 2. Participant Demographics

As experimental stimuli, using data from Open Food Facts (https://world.openfoodfacts.org/), five dairy-type food items (soy milk beverage, milk, yogurt, ice cream) ranging from relatively healthy to unhealthy were selected, one at each healthiness level. For each product, either a Nutri-Score label or an MTL label was displayed, and package images including the product name, flavor description, and product image were created (foods within the same category, selected based on the nutritional values of actual existing products). To ensure ecological validity, all package images were created by a professional package designer. To control for potential effects of visual salience, the size of the

area of interest (AOI) for the nutrition label was held constant across all stimuli (W747 × H378 pixels).

Insert Figure 1 here

Figure 1. Example of Experimental Stimuli (Product Packages with Nutri-Score and MTL Labels)

In terms of procedure, participants were first given an explanation of the ethics review procedures. Because neither label is currently used in Japan, printed materials describing both systems were provided. Subsequently, while participants viewed the stimuli on a computer monitor, eye movements were recorded using a Tobii Pro Fusion eye tracker. The procedure consisted of (1) gaze calibration, (2) task instructions asking participants to evaluate each product as if shopping and to press the Enter key when ready to evaluate purchase intention (5-point scale), (3) a practice trial, and (4) viewing of the five experimental stimuli followed by a post-viewing questionnaire survey. To eliminate order effects, the presentation order of the five stimuli was randomized for each participant.

In the data analysis, the primary dependent variables for the nutrition-label AOI were (1) time to first fixation (the elapsed time from stimulus onset until the first gaze landed on the AOI) and (2) total fixation count (the total number of fixations within each region). Because the sample size was relatively small and deviations from normality were observed (Shapiro–Wilk test), nonparametric statistical tests (Mann–Whitney U test) were adopted. To control the accumulation of Type I error from repeating tests across the five food categories, a Bonferroni correction (adjusted significance level $\alpha = .01$) was applied. In addition, given the

constraints of sample size, the effect size r was computed to examine the robustness of the phenomena.

Results

Time to First Fixation (Test of Hypothesis 1)

To test Hypothesis 1, the Mann–Whitney U test was conducted. For time to first fixation, even after applying the Bonferroni correction ($\alpha = .01$), a significant difference was retained only at healthiness level A ($p = .009$, $r = .48$). This suggests that the visual advantage of Nutri-Score is not uniformly exerted across all foods but becomes salient under specific conditions, especially when a "desirable evaluation (A)" is presented. Hypothesis 1 was therefore only partially supported.

Insert Table 3 here

Table 3. Comparison of Time to First Fixation Between Nutri-Score and MTL by Food Healthiness Level

As to why a significant difference was observed only at level A, in the case of healthy foods, Nutri-Score A clearly conveys the signal "this is a desirable food" through a single symbol, so it may have been preferentially attended to as a highly relevant source of information for the participants' evaluation task ("how much should one eat this?"). On the other hand, for less healthy foods, the evaluative usefulness of the information conveyed by the label becomes less

dependent on the symbolic form, and the difference between the two labels is thought to have narrowed.

Total Fixation Count: Heatmaps and Statistical Tests (Test of Hypothesis 2)

To examine overall patterns of visual attention, heatmaps aggregating total fixation counts across participants were generated. Across all products, participants commonly fixated on both the nutrition label and the flavor description. The trends observed in the heatmaps were also corroborated by the statistical tests reported below (Table 4).

Insert Figure 2 here

Figure 2. Heatmaps of Total Fixation Count for Nutri-Score and MTL Labels

Next, total fixation count on the nutrition-label AOI was compared between Nutri-Score and MTL using the Mann–Whitney U test. The analysis revealed significant differences meeting the Bonferroni-corrected significance level ($\alpha = .01$) in all categories except healthiness level B (A, C, D, E). In every category, the Nutri-Score group had fewer total fixations than the MTL group.

In terms of effect sizes, the intermediate-healthiness (C) and moderately unhealthy (D) foods yielded $r = .69$ and .66, substantially exceeding the criterion for a large effect ($r \geq .50$), and the unhealthy food (E) likewise showed a large effect of $r = .53$. The healthy food (A) showed $r = .48$, just short of the large-effect criterion but corresponding to a medium-to-large effect.

Even for condition B, where no significant difference was observed, a medium effect size of $r = .30$ was found, suggesting that a between-group difference itself exists, but that statistical power may have been reduced by the

large individual differences in the MTL condition (IQR = 19). In sum, Hypothesis 2 was supported in all categories except B.

Insert Table 4 here

Table 4. Comparison of Total Fixation Count for Nutri-Score and MTL by Food Healthiness Level

Examining the data for moderately healthy foods (Nutri-Score B) in detail, the interquartile range for MTL was extremely large at 19 (IQR for Nutri-Score B: 8), and variability in individual differences may have hindered the detection of a significant difference. The large individual differences in the MTL condition themselves are an interesting finding that suggests MTL processing may depend on individual consumers' involvement and nutritional literacy, and further investigation through replication is warranted.

Conclusion and Discussion

The central finding of this study is that Nutri-Score elicited significantly fewer total fixations than MTL, and that this difference was stably observed across four categories of differing healthiness (A, C, D, E; r = .48–.69). Hypothesis 2 was supported in all categories except B, making it the most robust finding of the study as a whole. The result can be said to support the findings of van Herpen & van Trijp (2011) and Gabor et al. (2020) in the Japanese consumer context, in which empirical evidence had been scarce.

This difference can be organized within the framework of Petty & Cacioppo's (1986) Elaboration Likelihood Model (ELM). Because Nutri-Score aggregates multiple pieces of nutritional information into a single evaluative symbol, consumers can arrive at a healthiness judgment merely by referring to a peripheral cue. MTL, in contrast, requires elaborative processing in which color information for each nutrient is individually referenced and integrated. The systematic difference in total fixation counts empirically supports—through visual-attention indices—the view that, although the two labels belong to the same FOPNL category, the modes of processing they elicit correspond respectively to peripheral-route and more central-route processing.

However, in the moderately healthy food (B) condition under MTL, individual differences in total fixation count were extremely large (IQR = 19). Although this is a post hoc observation and definitive conclusions must be reserved, it suggests that MTL may function for consumers with high nutritional literacy and high involvement in food choice, while it may not be fully utilized by consumers with low involvement. The perspective that the effect of MTL may depend on consumer characteristics is a consideration in evaluating the merits of uniform implementation, and it constitutes an issue for future investigation.

On the other hand, Hypothesis 1 regarding time to first fixation yielded the limited result of being supported only for healthy foods (A). Summary-evaluation labels

in general are not always detected more quickly; detection may be facilitated when a "desirable" symbol is used.

Limitations and Future Research

This study has several limitations. First, the sample size is small (N = 29) and unbalanced across groups (Nutri-Score: 12; MTL: 17). The fact that no significant differences were observed in Hypothesis 1 outside of healthiness level A, and that no significant difference was observed in Hypothesis 2 under condition B, is not evidence that effects do not exist but may instead reflect insufficient statistical power. Replications with larger samples are needed in future research.

Second, participants were limited to university students. Because interest in nutrition and purchasing experience differ across generations, verification with a more diverse range of participants is needed.

References

Andrews, J. C., Burton, S., & Kees, J. (2011). Is simpler always better? Consumer evaluations of front-of-package nutrition symbols. *Journal of Public Policy & Marketing, 30*(2), 175–190. https://doi.org/10.1509/jppm.30.2.175

Bialkova, S., Grunert, K. G., & van Trijp, H. C. M. (2013). Standing out in the crowd: The effect of information clutter on consumer attention for front-of-pack

nutrition labels. Food Policy, 41, 65–74. https://doi.org/10.1016/j.foodpol.2013.04.010

Egnell, M., Talati, Z., Hercberg, S., Pettigrew, S., & Julia, C. (2018). Objective understanding of front-of-package nutrition labels: An international comparative experimental study across 12 countries. *Nutrients, 10*(10), 1542. https://doi.org/10.3390/nu10101542

Fenko, A., Nicolaas, I., & Galetzka, M. (2018). Does attention to health labels predict a healthy food choice? An eye-tracking study. *Food Quality and Preference, 69*, 57–65. https://doi.org/10.1016/j.foodqual.2018.05.012

Gabor, A. M., Stojnić, B., & Ostić, D. B. (2020). Effects of different nutrition labels on visual attention and accuracy of nutritional quality perception–Results of an experimental eye-tracking study. *Food Quality and Preference, 84*, 103948. https://doi.org/10.1016/j.foodqual.2020.103948

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. Advances in Experimental Social Psychology, 19, 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2

Roberto, C. A., & Khandpur, N. (2014). Improving the design of nutrition labels to

promote healthier food choices and reasonable portion sizes. International *Journal of Obesity, 38*(Suppl 1), S25–S33. https://doi.org/10.1038/ijo.2014.86

Siegrist, M., Leins-Hess, R., & Keller, C. (2015). Which front-of-pack nutrition label is the most efficient one? The results of an eye-tracker study. Food Quality and Preference, 39, 183–190. https://doi.org/10.1016/j.foodqual.2014.07.010

van Herpen, E., & van Trijp, H. C. M. (2011). Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints, *Appetite, 57*(1), 148-160. https://doi.org/10.1016/j.appet.2011.04.011

Tables and Figures

Table 1. Prior Research on the Comparison of Nutrition Labels: Based on Eye-Tracking Experiments

Author	Compared Labels					Participants	Results
	Nutri-Score	Logo Label	MTL	GDA	Nutrition Facts		
van Herpen & van Trijp (2011)		○	○		○	Experiment 1: 309 university students in the Netherlands (Mean age 20.8)	・Nutritional fact panels (or Nutritional components) are less noticed compared to logo labels or MTLs (Multiple Traffic Lights), and their effectiveness in promoting healthy food choices is low. ・Logo labels are more likely to be noticed than other labels. This tendency is strongest when there is a specific health goal (such as a low-sodium target). ・Logo labels require a short fixation time (or short gaze duration).
						Experiment 2: 261 university students in Turkey (Mean age not specified)	・MTLs (Multiple Traffic Lights) and logo labels promote healthy food choices. ・The logo label, in particular, was effective even when consumers were under time pressure (or time constraints).
Bialkova et al. (2013)		○		○ (Colored / Monochrome))		Experiment 1: 80 university students (Country of residence, Mean age not specified)	・The colored GDA showed longer fixation time and more fixations than the monochrome GDA or logo labels. ・This tendency (the above pattern) was more pronounced for low-fat foods and when the consumer was health-conscious.
						Experiment 2: 80 university students (Different from the participants in Experiment 1)	・For health-conscious consumers, the colored GDA showed longer fixation time and more fixations than the monochrome GDA or logo labels. ・For low-fat foods, the colored GDA also showed longer fixation time and more fixations than the monochrome GDA or logo labels.
Siegrist et al. (2015)			○	○	○	98 people (Mean age 35) (Country of residence not specified)	・MTL (Multiple Traffic Light) was effective despite requiring a small amount of visual attention (or a low level of visual processing).
Fenko et al. (2018)		○	○			48 university cafeteria users in the Netherlands (Mean age 31)	* Compared to logo labels, MTL (Multiple Traffic Light) had a higher number of fixations and longer fixation time. * Participants with a high general interest in health focused their attention on the nutrition label under non-time-constrained conditions. Conversely, participants with a low general interest in health did not show this pattern.
Gabor et al. (2020)	○		○	○		76 adults in Croatia (Mean age 22) (including 56 nutritionists for comparison)	・Nutri-Score had a short total fixation time, a short mean fixation time, and a small number of fixations.

Note: The GDA (Guideline Daily Amounts) label indicates the amount of each nutrient per 100g of food, as well as the recommended daily intake for that nutrient.

Table 2. Participant Demographics

Group	N	Gender (M/F)	Age (Mean ± SD)
Nutri-Score	12	M6 /F6	20.8 ±1.2
MTL	17	M8 /F9	21.1 ±1.4

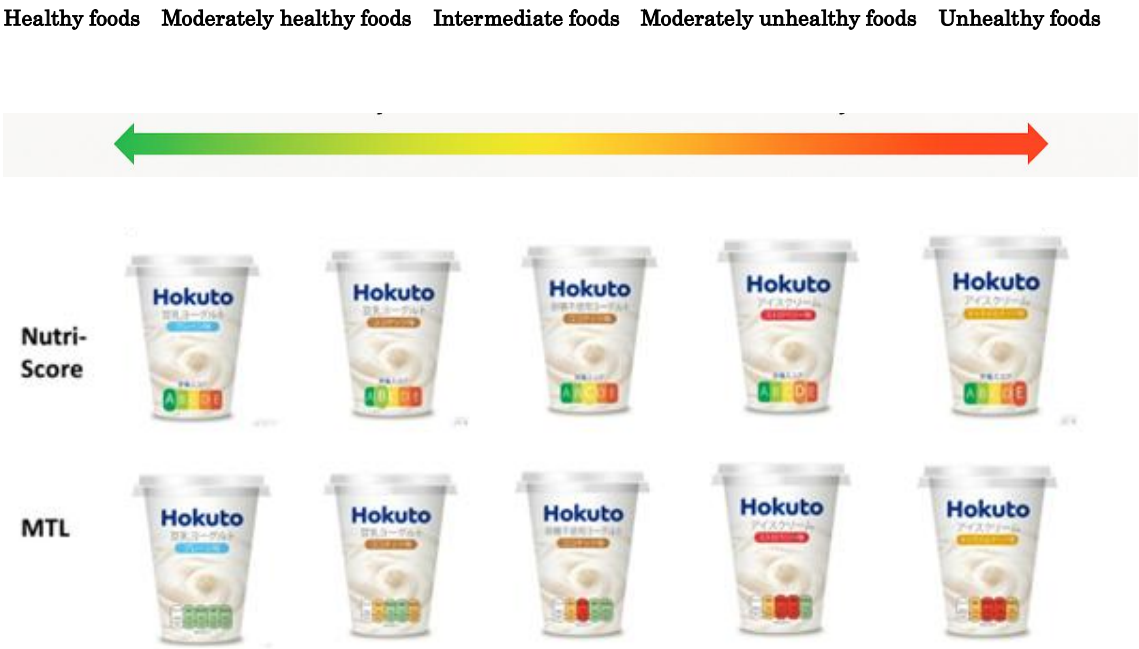

Figure 1: Example of Experimental Stimuli (Product Package with Nutri-Score and MTL Labels)

Table 3. Comparison of time to first fixation on Nutri-Score and Multiple Traffic Light (MTL) labels using the Mann–Whitney U test across five levels of food healthfulness.

Food healthfulness	NP model	*N*	Median (ms)	IQR	*U*	*r*	*p*	Result (α = .01)
Healthy foods	Nutri-Score (A)	12	803.5	1213	44.00	0.48	0.009	Significant
	MTL	17	1975.0	1587				
Moderately healthy foods	Nutri-Score (B)	12	1715.0	1032	82.00	0.16	0.394	n.s.
	MTL	17	2073.0	2113				
Intermediate foods	Nutri-Score (C)	12	1193.0	795	70.00	0.26	0.166	n.s.
	MTL	17	1896.0	1573				
Moderately unhealthy foods	Nutri-Score (D)	12	1303.5	1378	85.00	0.13	0.471	n.s.
	MTL	17	1471.0	1414				
Unhealthy foods	Nutri-Score (E)	12	1133.0	2184	79.00	0.18	0.325	n.s.
	MTL	17	1612.0	2044				

Note. NP = Nutrient Profiling; MTL = Multiple Traffic Light label; IQR = interquartile range. Time to first fixation is reported in milliseconds. Medians and IQRs are reported for each group. The effect size was calculated as $r = |Z| / \sqrt{N}$. Following the Bonferroni correction for five comparisons, the adjusted significance level was set at α = .01. The underlined p-value indicates a statistically significant difference; significant differences are shown in red.

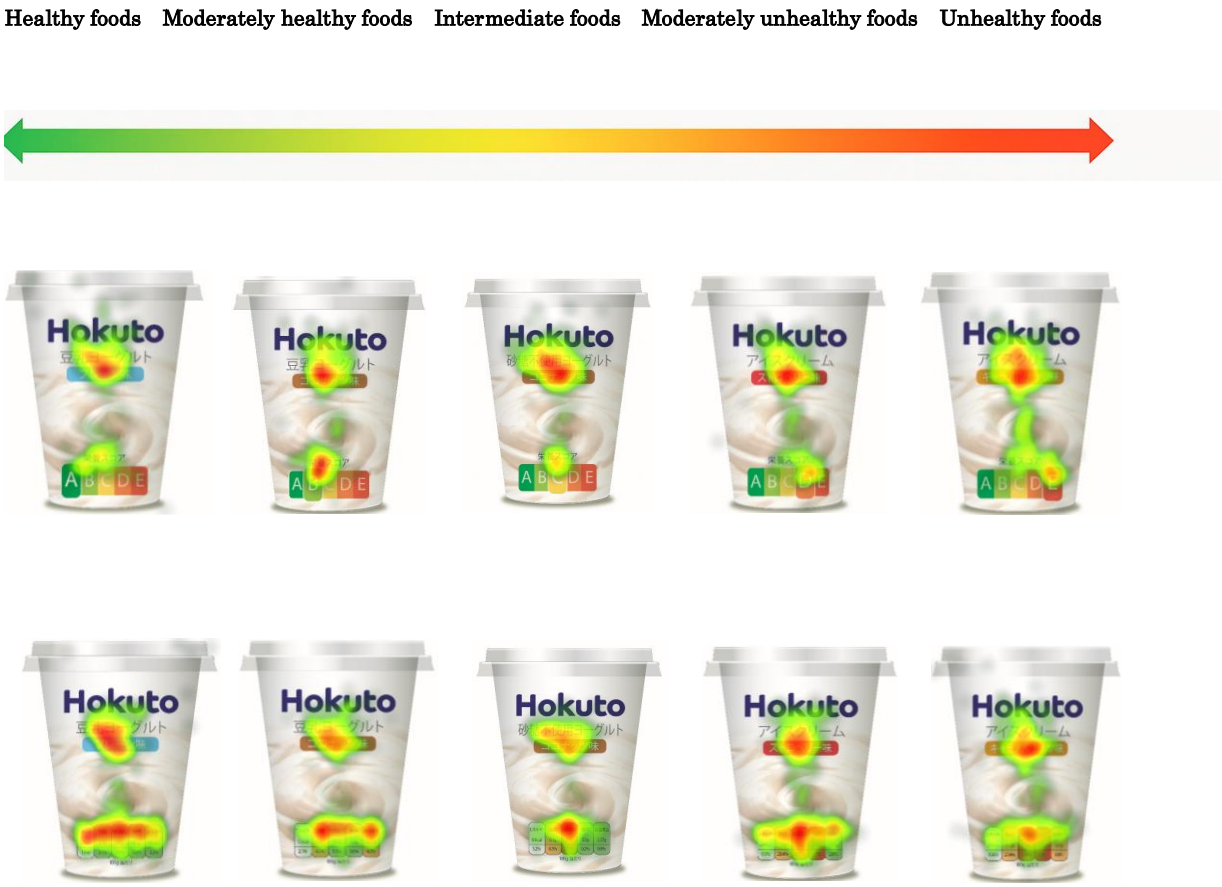

Figure 2: Heatmap of Total Number of Fixations for Nutri-Score Labels and MTL Labels

Table 4. Comparison of the total number of fixations on Nutri-Score and Multiple Traffic Light (MTL) labels using the Mann–Whitney U test across five levels of food healthfulness.

Food healthfulness	NP model	*N*	Median	IQR	*U*	*r*	*p*	Result (α = .01)
Healthy foods	Nutri-Score (A)	12	5.50	6	44.000	0.480	0.009	**Significant**
	MTL	17	16.00	14				
Moderately healthy foods	Nutri-Score (B)	12	6.00	8	65.500	0.300	0.107	n.s.
	MTL	17	16.00	19				
Intermediate foods	Nutri-Score (C)	12	4.00	5	19.500	0.690	< .001	**Significant**
	MTL	17	14.00	13				
Moderately unhealthy foods	Nutri-Score (D)	12	5.00	5	23.000	0.660	< .001	**Significant**
	MTL	17	13.00	8				
Unhealthy foods	Nutri-Score (E)	12	5.00	6	38.000	0.530	0.004	**Significant**
	MTL	17	10.00	12				

Note. NP = Nutrient Profiling; MTL = Multiple Traffic Light label; IQR = interquartile range. Medians and IQRs are reported for each group. The effect size was calculated as $r = |Z| / \sqrt{N}$. Following the Bonferroni correction for five comparisons, the adjusted significance level was set at α = .01. Significant differences are shown in red.