Visual Attention to Nutri-Score among Japanese Consumers: The Role of Food Health Image and Schema Incongruence

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

[a]Faculty of Information and Communication, Bunkyo University, Kanagawa, Japan; [b]Faculty of Health and Nutrition, Bunkyo University, Kanagawa, Japan

*Corresponding author: Masao Okano, okano@shonan.bunkyo.ac.jp

Masao Okano is a Professor at the Faculty of Information and Communication, Bunkyo University, Japan. His research interests include consumer behaviour, health communication, and data analysis. 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.

Visual Attention to Nutri-Score among Japanese Consumers: The Role of Food Health Image and Schema Incongruence

Abstract

Drawing on schema incongruence theory, this eye-tracking study examined whether Japanese consumers' visual attention to Nutri-Score varies with the congruence between label grade and prior health images of food categories. Fifty-two Japanese university students viewed fictitious packages from three categories differing in health image (yogurt, ice cream, cup noodles) displaying either Grade A (favorable) or Grade E (unfavorable). Linear mixed-effects models on total fixation duration, fixation count, and time to first fixation revealed a significant category × grade interaction: heightened attention emerged only when Grade E appeared on yogurt, a high-health-image food. Grade A on low-health-image foods produced no comparable effect, and time to first fixation did not differ across conditions. The incongruence effect was thus asymmetric, with negative labels on healthy foods acting as warning cues with high diagnostic value. FOPNL effects emerge through interaction with consumers' category schemata.

Keywords: eye-tracking; front-of-pack nutrition label; Nutri-Score; schema incongruence; health image; Japan

Introduction

Communicating nutritional information effectively to consumers remains a central challenge in food policy and marketing. In Japan, mandatory nutrition information is typically presented in tabular form on the back of food packages. Although nutrition tables provide

detailed information, prior research consistently shows that consumers rarely refer to them in real purchase situations, and their influence on food choice behavior is limited.

Front-of-pack nutrition labels (FOPNLs) have been introduced as a complementary system that summarizes nutritional attributes into simplified visual formats processable under limited time and attention. Among various systems, Nutri-Score has received particular attention (Julia & Hercberg, 2017; Egnell et al., 2018).

Accordingly, various studies have been conducted on the visibility, comprehension, and consumer-preference responses to Nutri-Score (Egnell et al., 2019; Vandevijvere et al., 2020; Hagmann & Siegrist, 2020; van den Akker et al., 2022). However, while prior eye-tracking research has identified bottom-up determinants of label attention—size, color, placement, and familiarity (Bialkova & Van Trijp, 2010; Bialkova et al., 2014; Visschers et al., 2010)—no studies have examined the influence of top-down factors on visibility, particularly the congruence between label information and consumers' prior product schemata.

Schema incongruence theory argues that expectation-violating information attracts greater attention and deeper processing (Heckler & Childers, 1992; Mandler, 1982; Meyers-Levy & Tybout, 1989), predicting that Nutri-Score functions most effectively as an attention cue when it contradicts intuitive beliefs about a product's healthfulness. Moreover, although Nutri-Score was developed and primarily validated in the European context, its absence from Japan's domestic labeling system makes Japan an important non-European test case in which familiarity-based processing cannot be presupposed.

We therefore set the following research question.

RQ: Does Japanese consumers' visual attention to Nutri-Score differ depending on the congruence between the label grade and the food product's prior health image?

Theoretical Background

Schema incongruence theory posits that individuals possess cognitive schemata—organized

knowledge structures that guide perception and interpretation (Mandler, 1982). Congruent information is processed fluently and heuristically, whereas incongruent information generates arousal and motivates additional cognitive effort to resolve the discrepancy. Prior research has shown that moderate incongruence enhances attention, elaboration, and memory (Lee & Schumann, 2004; Meyers-Levy & Tybout, 1989).

Mandler (1982) proposed an inverted-U relationship between the degree of discrepancy and cognitive engagement: moderate incongruence—surprising yet interpretable—generates the greatest elaboration, whereas extreme incongruence may trigger rejection or disengagement rather than deeper processing.

In food choice contexts, consumers hold well-established schemata regarding the healthfulness of familiar food categories. When a nutrition label is consistent with these expectations, it provides little new information and may receive minimal attention. Conversely, when the label violates expectations, consumers may allocate greater attentional resources to resolve the inconsistency. Because food choices are typically made under low involvement and limited cognitive resources, peripheral cues that stand out visually or conceptually play a disproportionately large role in guiding attention (Van Herpen & Van Trijp, 2011). Nutri-Score, as a highly salient visual cue, is therefore well suited to trigger expectancy-based attention effects, but the magnitude of these effects is expected to depend on the category-specific degree of incongruence and consumers' perceived diagnosticity of the label. We therefore set the following hypotheses.

H1: Nutri-Score A attracts more visual attention when displayed on low-health-image foods than on high-health-image foods.

H2: Nutri-Score E attracts more visual attention when displayed on high-health-image foods than on low-health-image foods.

Method

Experimental Stimuli

Considering ecological validity, a professional package designer created fictitious food packages. Three food categories with clearly differing perceived health images among Japanese consumers were selected (yogurt, ice cream, and cup noodles). Yogurt is generally regarded as healthy, whereas ice cream and cup noodles are perceived as less healthy. With the exception of the product name and the displayed Nutri-Score (A or E), all packages were identical in layout, color scheme, typography, and visual elements, yielding six target stimuli (3 categories × 2 grades). To avoid presenting the three target categories consecutively, two filler products (chocolate and cola) were presented between them (in the second and fourth positions). In addition to the Nutri-Score, a tabular nutrition display was placed on the right side of each package. However, because the present study focused primarily on visual attention to the Nutri-Score, AOI analyses were restricted to the Nutri-Score region.

[Figure 1 near here]

Participants and Design

A total of 52 Japanese university students participated between November and December 2024 (41 female, 11 male; M age = 20.37 years, SD = 1.31; range 19–22; normal or corrected-to-normal vision; individual lab sessions). Because presenting both grades of the same food (e.g., yogurt A and yogurt E) to the same participant would be unnatural, each food category was shown in only one grade per participant (incomplete block design). As a result, the target stimuli were distributed such that each participant evaluated three food–grade combinations along with two fillers (five stimuli in total).

Stimulus assignment and presentation order were fully randomized.

Apparatus and Procedure

Eye movements were recorded using a Tobii Pro Fusion eye tracker (Tobii Technology,

Stockholm, Sweden; 250 Hz; gaze accuracy 0.3°; precision 0.2° RMS; latency < 13 ms; operating distance 50–80 cm; head movement tolerance 40 cm × 25 cm at 65 cm), displayed on a 24-inch monitor, and analyzed using Tobii Pro Lab. Participants first received a printed explanation of Nutri-Score. This procedure was necessary given the extremely limited familiarity with Nutri-Score in Japan. However, this prior exposure may have artificially heightened attention to the Nutri-Score; we position this as a limitation of the study. After a comprehension check and 9-point calibration, participants imagined a scenario of selecting foods for personal consumption and pressed the Enter key once they had decided how much to purchase each product. Following a practice trial, the five stimuli were presented one at a time in randomized order. Each stimulus was preceded by a 5-second fixation cross. A post-experiment questionnaire measured the perceived healthfulness of the foods presented.

Measures

Visual attention was assessed by three AOI measures: total fixation duration (ms), fixation count, and time to first fixation (ms). The AOI was the rectangular region encompassing the Nutri-Score graphic and was kept identical in size across all stimuli (83 × 169 pixels, width × height). Fixations were identified using the Tobii I-VT filter (velocity threshold 30°/s; minimum fixation duration 60 ms; merging of adjacent fixations within 75 ms and 0.5°). Perceived health image was measured on a 7-point rating scale (1 = very unhealthy, 7 = very healthy).

Results

Health Image Manipulation Check

A repeated-measures one-way ANOVA confirmed significant differences in perceived health image (sphericity met: $W = .996$, $p = .914$). The main effect of food category was significant, $F(2, 102) = 538.81$, $p < .001$, partial $\eta^2 = .914$. Yogurt received the highest ratings ($M = 6.38$,

SD = 0.60), followed by ice cream (M = 2.40, SD = 1.03) and cup noodles (M = 1.67, SD = 0.79). All Bonferroni-corrected pairwise comparisons were statistically significant (all $p < .001$).

[Figure 2 near here]

Visual Attention to Nutri-Score

Separate LMMs were estimated for each dependent variable, with food category and label grade as fixed effects and participant as a random effect. To examine the hypotheses, we first evaluated the overall interaction using likelihood ratio tests, followed by contrast analysis based on estimated marginal means (EMMs).

Total fixation duration: The overall interaction was significant, $\chi^2(2) = 6.92$, $p = .032$. To test H2 (negative incongruence), contrast analysis within the Grade E condition revealed that yogurt attracted significantly longer fixation durations than the average of the other categories ($p = .014$). In contrast, comparisons within the Grade A condition (testing H1) yielded no significant differences ($R^2m = .168$; $R^2c = .400$).

Fixation count: The interaction was significant, $\chi^2(2) = 10.11$, $p = .006$. Contrast analysis confirmed that yogurt received significantly more fixations in the Grade E condition compared with the other categories ($p = .002$; $R^2m = .173$; $R^2c = .381$).

Time to first fixation: Neither the main effects nor the interaction were significant ($p > .10$), indicating that initial attentional capture did not vary as a function of food category or label grade.

Across both total fixation duration and fixation count, heightened visual attention was observed exclusively for yogurt–Grade E. That is, H2 was supported, whereas H1 was not. The uniformity of time to first fixation across conditions indicates that between-condition differences reflect post-initial elaborative processing rather than differences in initial attentional capture. Descriptive statistics are presented in Table 1.

[Table 1 near here]

Heatmap Analysis

Heatmaps corroborated the statistical results, showing notably greater concentration of fixations on the Nutri-Score AOI in the yogurt–Grade E condition. No comparable concentration was observed for ice cream or cup noodles under either label grade.

[Figure 3 near here]

[Figure 4 near here]

Discussion

The present study examined whether Japanese consumers' visual attention to Nutri-Score depends on the congruence between label information and prior food health images. The results provided partial support for the hypotheses. H2 was supported only for yogurt: total fixation duration and fixation count were significantly higher for Grade E than for Grade A, consistent with the prediction that an unfavorable label on a high-health-image food triggers elaborative attention through schema incongruence. H1, in contrast, was not supported; for ice cream and cup noodles, Grade A did not attract significantly greater visual attention than Grade E. Furthermore, no significant effects emerged for time to first fixation, and initial attentional capture was uniform across all food–grade combinations, indicating that subsequent fixation differences reflect post-initial elaborative processing.

In interpreting these category-specific patterns, it is important to note first that the incongruence effects observed in this study were not symmetric. Whereas increased visual attention was confirmed when Grade E was displayed on yogurt—a food generally regarded as healthy—comparable increases were not observed when Grade A was displayed on ice cream and cup noodles, which are perceived as relatively unhealthy. This suggests that visual attention to Nutri-Score is not determined merely by the presence or absence of incongruence

with expectations, but may also depend on the direction of incongruence.

This asymmetry can be interpreted as reflecting the higher diagnostic value of negatively incongruent information. In particular, when an unfavorable label grade is assigned to a food considered healthy, the information not only contradicts existing expectations but is also likely to function as a warning cue that prompts reconsideration of the evaluation and choice of the product. Consequently, in the yogurt × Grade E condition, total fixation duration and fixation count are considered to have increased in a manner that reflects post-initial elaborative processing rather than initial attentional capture.

In contrast, the absence of significant differences when a favorable label grade was assigned to a low-health-image food suggests that positively incongruent information may not have elicited attention to the same degree as negatively incongruent information. In particular, because cup noodles had an especially low health image, even Grade A may not have been recognized as sufficiently credible or diagnostic, which may have prevented an increase in visual attention.

Limitations

Four limitations should be acknowledged. First, all participants received a printed Nutri-Score explanation before the experiment. Although this was necessary for meaningful label processing, it constitutes an artificial priming that may have heightened AOI fixation counts and amplified incongruence effects relative to naturalistic shopping contexts. Second, the sample of 52 university students limits generalizability, and a representative sample spanning a wider range of age, education, and dietary engagement is needed. Third, because the incomplete block design was structured so that the same participant did not view different grades of the same food category, statistical power may have been reduced relative to a fully within-subjects design. Fourth, visual attention to the tabular nutrition information juxtaposed within the stimuli, as well as the allocation of attention between this tabular information and

Nutri-Score, was not included in the present analysis. Future research is required to clarify how tabular nutrition information is viewed in the presence of Nutri-Score.

Notwithstanding these limitations, the present study is meaningful in that it demonstrated that visual attention to Nutri-Score may vary depending on its relationship with prior health images of food categories. In particular, the finding that visual attention increased when an unfavorable evaluation was assigned to a food generally regarded as healthy suggests that FOPNL effects do not arise uniformly upon label presentation but emerge through interaction with consumers' category schemata. This indicates the need to consider category-specific differences in the implementation and evaluation of FOPNLs.

Conclusion

The present study demonstrated that Japanese consumers' visual attention to Nutri-Score varies with the congruence between prior health images of food categories and the label grade. Specifically, total fixation duration and fixation count increased only when Grade E was displayed on yogurt, whereas no comparable increases were observed when Grade A was displayed on ice cream or cup noodles. Furthermore, because no significant differences were observed in time to first fixation, the observed differences are interpreted as reflecting post-initial elaborative processing of incongruent information rather than initial attentional capture.

These findings suggest that, even in countries and regions where consumer exposure to front-of-pack nutrition labels is relatively limited, FOPNL effects do not arise uniformly but are governed by their relationship with category schemata, including the direction of incongruence. Therefore, when examining the effects of Nutri-Score implementation, it is important to adopt an approach that considers not only differences in health images across food categories but also whether the resulting incongruence is positive or negative in direction.

References

Bialkova, S., & Van Trijp, H. C. M. (2010). What determines consumer attention to nutrition labels? *Food Quality and Preference, 21*(8), 1042–1051. https://doi.org/10.1016/j.foodqual.2010.07.001

Bialkova, S., Grunert, K. G., Juhl, H. J., Wasowicz-Kirylo, G., Stysko-Kunkowska, M., & Van Trijp, H. C. M. (2014). Attention mediates the effect of nutrition label information on consumers' choice: Evidence from a choice experiment involving eye-tracking. *Appetite, 76*, 66–75. https://doi.org/10.1016/j.appet.2013.11.021

Egnell, M., Talati, Z., Hercberg, S., Pettigrew, S., & Julia, C. (2018). Objective understanding of the Nutri-Score front-of-pack label by European consumers and its effect on food choices: An online experimental study. *Nutrients, 10*(10), 1542. https://doi.org/10.3390/nu10101542

Egnell, M., Talati, Z., Pettigrew, S., Galan, P., Hercberg, S., & Julia, C. (2019). Comparison of front-of-pack labels to help German consumers understand the nutritional quality of food products. *Ernahrungs Umschau, 66*(5), 76–84. https://doi.org/10.4455/eu.2019.020

Hagmann, D., & Siegrist, M. (2020). Nutri-Score, multiple traffic light and incomplete nutrition labelling on food packages: Effects on consumers' accuracy in identifying healthier snack options. *Food Quality and Preference, 83*, 103894. https://doi.org/10.1016/j.foodqual.2019.103894

Heckler, S. E., & Childers, T. L. (1992). The role of expectancy and relevancy in memory for verbal and visual information: What is incongruency? *Journal of Consumer Research, 18*(4), 475–492. https://doi.org/10.1086/209275

Julia, C., & Hercberg, S. (2017). Development of a new front-of-pack nutrition label in France: The five-colour Nutri-Score. *Public Health Panorama, 3*(4), 712–725.

Lee, E.-J., & Schumann, D. W. (2004). Explaining the special case of incongruity in advertising: Combining classic theoretical approaches. *Marketing Theory, 4*(1–2), 59–90. https://doi.org/10.1177/1470593104044087

Mandler, G. (1982). The structure of value: Accounting for taste. In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition: The 17th annual Carnegie symposium on cognition* (pp. 3–36). Lawrence Erlbaum Associates.

Meyers-Levy, J., & Tybout, A. M. (1989). Schema congruity as a basis for product evaluation. *Journal of Consumer Research, 16*(1), 39–54. https://doi.org/10.1086/209192

van den Akker, K., Bartelet, D., Brouwer, L., Luijpers, S., Nap, T., & Havermans, R. (2022). The impact of the Nutri-Score on food choice: A choice experiment in a Dutch supermarket. *Appetite, 168*, 105664. https://doi.org/10.1016/j.appet.2021.105664

Vandevijvere, S., Vermote, M., Egnell, M., Galan, P., Talati, Z., Pettigrew, S., Hercberg, S., & Julia, C. (2020). Consumers' food choices, understanding and perceptions in response to different front-of-pack nutrition labelling systems in Belgium: Results from an online experimental study. *Archives of Public Health, 78*, 30. https://doi.org/10.1186/s13690-020-00404-3

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

Visschers, V. H. M., Hess, R., & Siegrist, M. (2010). Health motivation and product design determine consumers' visual attention to nutrition information on food products. *Public Health Nutrition, 13*(7), 1099–1106. https://doi.org/10.1017/S1368980009993235

Table 1. *Descriptive Statistics (M and SD) for Total Fixation Duration (ms), Fixation Count, and Time to First Fixation (ms) by Food Category and Nutri-Score Grade*

Food Category	Grade	Total Fix. Dur. (M)	SD	Fix. Count (M)	SD	Time 1st Fix. (M)	SD
Ice cream	A	1677.6	1093.4	5.0	3.3	1723.3	1632.7
	E	1402.2	769.6	3.8	1.9	1882.0	1187.9
Yogurt	A	1280.1	945.4	3.9	2.0	1748.1	1624.6
	E	1739.2	1292.9	5.8	3.8	2368.6	1912.5
Cup noodles	A	1258.6	833.3	3.9	1.8	1422.0	635.8
	E	1119.7	742.6	3.8	2.2	2130.4	2277.5

Figure 1. Experimental stimuli used in the eye-tracking experiment.



Figure 2. Mean health image ratings by food category.

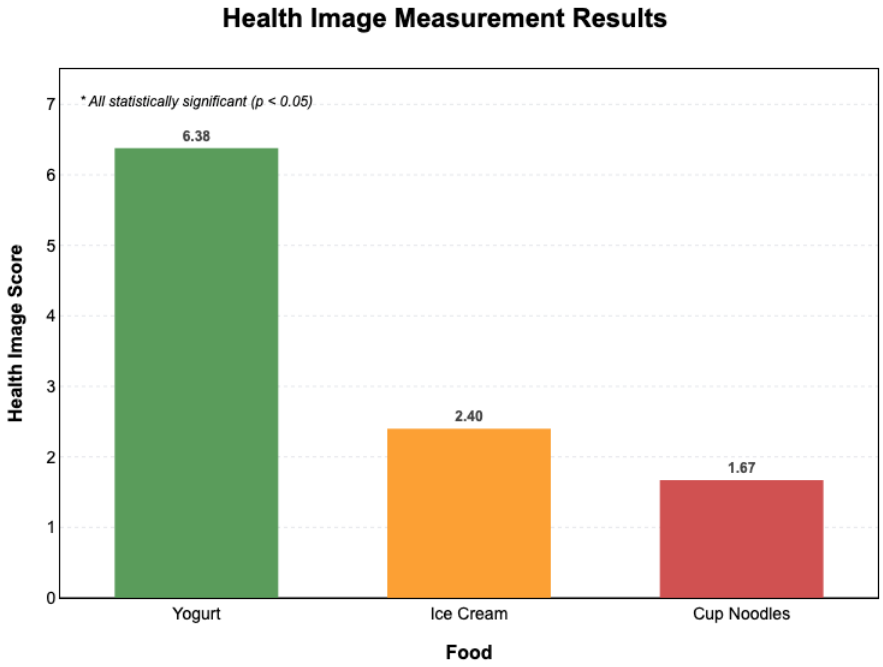

Figure 3. Fixation heatmaps for Nutri-Score A.



Figure 4. Fixation heatmaps for Nutri-Score E.

