Capturing the Healthfulness of the In-store Environments of United Kingdom Supermarket Stores Over 5 Months (January−May 2019)

Georgina Harmer, MSc,1 Susan A. Jebb, PhD,1 Georgia Ntani, PhD,2 Christina Vogel, PhD,2 Carmen Piernas, PhD1

Introduction: Numerous environmental factors within supermarkets can influence the healthfulness of food purchases. This research aims to identify the changes in store healthfulness scores and assess the variations by store type and neighborhood deprivation using an adapted Consumer Nutrition Environment tool.

Methods: Between January and May 2019, a total of 104 supermarkets in London were surveyed on 1−3 occasions. The adapted Consumer Nutrition Environment tool included data on 9 variables (variety, price, quality, promotions, shelf placement, store placement, nutrition information, healthier alternatives, and single fruit sale) for 11 healthy and 5 less healthy food items. An algorithm was used to create a composite score of in-store healthfulness and to assess inter-rater reliability. Longitudinal changes in overall store healthfulness and individual variables were investigated using multivariable hierarchical mixed models. Descriptive statistics were used to describe the differences by store type and neighborhood deprivation in each month. All analyses were conducted between January and July 2020.

Results: The adapted Consumer Nutrition Environment tool showed acceptable inter-rater reliability. Large stores exhibited healthier environments than small stores (p<0.001), with a similar pattern for each of the 9 individual variables. Within large stores, the overall healthfulness score did not change over the study period. Promotions on more healthful items increased in February (p=0.04), and the availability of healthier alternatives for less healthy foods decreased in March (p=0.01). Within small stores, there was a trend toward increasing healthfulness (p<0.001), primarily owing to more promotions on healthy items (p<0.001). There was no difference in overall healthfulness by neighborhood deprivation.

Conclusions: The adapted Consumer Nutrition Environment tool is sensitive to longitudinal changes in environmental variables that contribute to store healthfulness. A wider application of this tool could be used to map in-store environments to identify targets for interventions to encourage healthier food purchasing.

INTRODUCTION

Supermarkets are a major source of food purchased for home consumption, accounting for approximately 87% of all United Kingdom (UK) retail grocery sales.1 The supermarket food environment, including variables that influence purchasing such as availability, price, promotions, placement, variety,
quality, and nutrition information, has been found to influence dietary behaviors by encouraging both excessive food purchasing and sales of less healthy items.\(^2\)\(^-\)\(^4\) Understanding how key variables and the overall healthfulness of supermarket environments change over time is needed to identify and monitor the interventions most likely to be effective in encouraging healthier food purchasing.

Tools have been developed to assess supermarket environments, but they differ in their content and complexity, and their use is mostly limited to cross-sectional studies.\(^5\)\(^,\)\(^6\) The Consumer Nutrition Environment (CNE) tool was originally developed to capture information on products used to characterize healthy or less healthy dietary patterns of women of childbearing age.\(^7\) Adapting this tool to be suitable for the broader UK population will increase its applicability and use in evaluating supermarket interventions or policies. To the authors’ knowledge, there are no studies investigating the changes in the overall in-store environment of supermarkets over time. Yet, the supermarket environment is prone to rapid changes in promotional strategies that influence short-term purchasing behavior.\(^8\)

This study aims to describe the development and application of an adapted CNE tool and investigate changes in the healthfulness scores and the 9 contributing variables over the first 5 months of the year and to assess variations by store type and neighborhood deprivation. This period of the year is marked by numerous occasions typically celebrated with confectioneries, including Valentine’s Day, Mother’s Day, and Easter, which enables the assessment of seasonal fluctuations in supermarket environments.

**METHODS**

The CNE tool was originally developed to measure 9 variables (variety, price, quality, promotions, shelf placement, store placement, nutrition information, healthier alternatives, and single fruit sale) each assessed for 7 healthy and 5 less healthy foods.\(^7\)\(^,\)\(^9\)\(^,\)\(^12\) The food items were expanded in this study to better reflect the diet of the whole population. A greater range of fruits and vegetables frequently consumed in the UK and frequently consumed less healthy foods that represent target products in the UK government’s calorie reduction and sugar reduction programs were included.\(^3\)\(^,\)\(^4\)\(^,\)\(^14\) The adapted tool included 11 healthy (peppers, tomatoes, apples, bananas, carrots, broccoli, cauliflower, tinned tomatoes, tinned sweet corn, frozen mixed vegetables, and frozen peas) and 7 less healthy (sausages, crisps, carbonated soft drinks, biscuits, sweets, chocolate bars, and pizza) items. Only fruits and vegetables were included as healthy items because consumption of these products remains well below the recommended levels in the UK and selections can vary across store types and area deprivation.\(^15\) The adapted CNE tool and protocol are available in Appendix Files 2 and 3 (available online).

**Study Sample**

Stores were sampled across London from 2 major UK food retail chains, which jointly have >40% of the UK market share.\(^16\) Large and small stores were selected to cover a range of socioeconomic deprivation strata on the basis of the 2019 English Index of Multiple Deprivation (IMD) income domain, the official measure of relative deprivation in small areas in England (Table 1).\(^5\)\(^,\)\(^17\)\(^-\)\(^20\) The sample covered neighborhoods from Deciles 1–8, regrouped into IMDs 1–3 (high), 4–5 (middle), and 6–8 (low) (Table 2). The sample included 88% of the 2 target retailers within 2 highly deprived London areas, plus an additional 22 stores to increase representation across area deprivation levels.

**Measures**

The variables measured and methods used in this study were the same as those for the original CNE tool.\(^7\) Briefly, for each product, the number of varieties, price, promotion, shelf placement, and store placement were collected. Price was recorded for the cheapest item for each product and converted to pounds per portion. The type of nutrition information present and availability of a healthier alternative were collected only for less healthy products. The quality and option for single sale of 2 fruits and the quality of 5 vegetables were also assessed. Appendix Table 1 (available online) provides the full definitions and measurement scales of the variables assessed. Appendix 1 (available online) provides information on healthfulness score development.

Between January and May 2019, trained fieldworkers visited stores on 1–3 occasions in different months and completed the adapted CNE survey in paper format. For each time point, fieldworkers had a 2-week window in which to visit stores. Unexpected store manager refusal on 7 occasions and fieldworker unreliability meant that the final data set did not include 3 visits for all stores. A total of 8 stores were additionally surveyed by the primary researcher (GH) to assess the inter-rater reliability. All data were cleaned by GH and transferred to an electronic database before analysis.

**Statistical Analysis**

For inter-rater reliability analyses, the \(\kappa\) statistic was calculated to assess the level of agreement for each variable, except for price. The relative consistency of price responses was assessed using the coefficient of variation: the SD of the difference divided by the mean, expressed as a percentage (%).

**Table 1. Retail Food Outlet Categorization System**\(^9\)\(^,\)\(^17\)

<table>
<thead>
<tr>
<th>Code</th>
<th>Store type</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Large supermarket</td>
<td>≥5 manned cash registers, All foods and many varieties, Majority of supermarket share</td>
<td>Tesco Superstore, Sainsbury’s Superstore</td>
</tr>
<tr>
<td>2</td>
<td>Small supermarket</td>
<td>1–4 manned cash registers, Smaller store of known brand name</td>
<td>Tesco Express, Sainsbury’s Local</td>
</tr>
</tbody>
</table>

www.ajpmonline.org
Stores measured at least twice were included in the longitudinal analysis. Hierarchical mixed models with random intercepts for store, with an interaction term for store type by time and adjusted for deprivation level and retailer, were used. Postestimation marginal mean scores were calculated for each month and store type. A test for trend was performed with the same models described earlier, except that the coefficient for month was now modeled as a continuous variable.

Cross-sectional differences by store type and neighborhood deprivation for each month were illustrated using box plots, Student’s t-tests for normally distributed variables, Mann–Whitney U tests for nonparametric variables, and chi-square tests for categorical variables.

Exploratory analyses investigated whether the changes observed in individual variable scores were driven by changes in healthy or less healthy items. Hierarchical mixed models were used as described earlier, where the outcomes were the average scores from the healthy and less healthy items for price, variety, shelf placement, store placement, and promotions.

A sensitivity analysis was performed to recalculate overall in-store healthfulness scores after subtracting 2 variables (quality and single fruit sale) that showed poor inter-rater reliability or little variation across time or store type, respectively. Overall healthfulness scores were also recalculated after reducing the number of healthy items to 7 to match the number of included less healthy items. Fresh tomatoes, cauliflower, tinned sweet corn, and frozen peas were removed to leave the 2 original fruits (apples and bananas) and the other most commonly consumed vegetables in the UK (peppers, broccoli, carrots, tinned tomatoes, and frozen mixed vegetables).

Descriptive statistics and regressions were conducted using Stata, version 14. A 2-sided p<0.05 was used to define statistical significance.

RESULTS

Data were collected from a total of 104 stores (Table 2), mostly small supermarkets (85%, n=88), and from the most deprived areas with IMDs 1–3 (n=47, 46%).

Overall inter-rater reliability of the tool was good for most components except for shelf and store placement and quality of fruits and vegetables (Appendix 2, available online).

Longitudinal analyses showed that at all timepoints, large stores had a higher overall healthfulness score (more healthful environments) and higher scores for variety, promotions, shelf placement, store placement, quality, healthier alternative availability, and nutrition information than small stores (Figure 1). It was only price that had a higher score in small stores than in large stores, reflecting prices more supportive of affordable healthy food purchasing than less healthy foods.

For large stores, there was no overall trend and no significant changes in the overall healthfulness score in any month compared with that in January. Among individual variables, the score for promotions significantly increased in February (p=0.04), reflecting a healthier promotional environment, decreased in March before increasing again in May, with no significant trend across the study period. Availability of healthier alternatives to less healthy foods declined in March (p=0.01).

For small stores, there was a significant decrease in the overall healthfulness score in February compared with that in January (p=0.03) but an overall significant upward trend from January to May (p<0.001) (Figure 1 and Appendix Table 3, available online). In March, price scores decreased (p=0.03), and less healthful items became cheaper; however, there was no overall trend. There was a trend toward increasing healthfulness of promotions between January and May (p<0.001), and by May, this was significantly different from that in January (p<0.01), reflecting a healthier promotional environment with time.

Changes in healthy and less healthy items separately (Figure 2 and Appendix Table 4, available online) showed that in January–May, in both large and small stores, less healthy items were cheaper and more frequently promoted than healthier items. The price difference between less healthy and healthy items was greater in small stores than in large stores (Figure 2). In large stores, the price of the cheapest healthy (p=0.01) and less healthy (p=0.02) items increased over time. This was also observed in small stores for healthy (p<0.001) but not for less healthy items. In large stores, there was a greater variety of healthy items than less healthy items across the whole study period, whereas in small stores, there was a similar variety of less healthy items to that in large stores but far fewer varieties of healthy items. Shelf and store placement scores highlighted that healthy items tend to be more

<table>
<thead>
<tr>
<th>Store type</th>
<th>IMDs 1–3 (most deprived)</th>
<th>IMDs 4–5</th>
<th>IMDs 6–8 (least deprived)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large stores</td>
<td>Retailer 1</td>
<td>Retailer 2</td>
<td>Retailer 1</td>
</tr>
<tr>
<td>Small stores</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Total, n (%)</td>
<td>32 (34)</td>
<td>12 (12)</td>
<td>24 (23)</td>
</tr>
</tbody>
</table>

IMD, Index of Multiple Deprivation.
Figure 1. Changes in store healthfulness score and individual variables across January–May 2019 by store type. Note: Boldface indicates statistical significance (*p<0.05; **p<0.01; ***p<0.001). Estimates from hierarchical mixed models with random intercepts for store and retailer, the interaction of store type X time; and adjusted for IMD score; $P_{\text{trend}}$ between January and May are shown next to the line plots for each store type separately. Small stores: January $n=28$, February $n=49$, March $n=37$, May $n=56$. Large stores: January $n=12$, February $n=10$, March $n=7$, May $n=15$. IMD, Index of Multiple Deprivation; $P_{\text{trend}}$, p-values for the trend.
Figure 2. Changes in individual variables of the store healthfulness score for healthy and less healthy products compared with that in January in (A) large stores and (B) small stores.

Note: Boldface indicates statistical significance (*p<0.05; **p<0.01; ***p<0.001). Estimates from hierarchical mixed models with random intercepts for store and retailer, the interaction of store type X time; and adjusted for IMD score; P_trend between January and May are shown next to the line plots for each store type separately. Small stores: January n=28, February n=49, March n=37, May n=56. Large stores: January n=12, February n=10, March n=7, May n=15.

IMD, Index of Multiple Deprivation; P_trend, p-values for the trend.
easily identifiable than less healthy items in both store types every month (Figure 2).

In large stores between January and February, promotions of healthier items increased more than promotions of less healthy items ($p<0.001$ and $p=0.01$, respectively). Furthermore, there was an increasing trend of promotions on healthy items across January–May ($p<0.01$ for trend). In February, prominence in shelf placement of healthy items ($p=0.02$) was reduced compared with that in January, with little change in the prominence of less healthy items.

In small stores, there was a significant fall in the price of the cheapest less healthy items ($p=0.01$) in March compared with that in January. In February and May, there were significant increases in the promotions on healthy items ($p<0.01$ and $p<0.001$, respectively) compared with that in January and an overall trend of increasing the promotions on healthy items across January–May ($p<0.01$ for trend). Store prominence of less healthy items increased significantly in February ($p=0.01$) from January, whereas store prominence of healthy items deteriorated.

Cross-sectional analyses showed that healthfulness scores for small stores were lower than in large stores at all timepoints ($p<0.001$) (Appendix Figure 1A, available online). The individual variables followed similar patterns, with significantly higher estimates at all timepoints in large stores except for price scores, which were significantly higher in small stores (Appendix Table 5, available online).

There was no clear pattern between the composite healthfulness score or the individual variables and neighborhood deprivation across all time points (Appendix Table 3, available online), nor at any given timepoint (Appendix Table 6, available online, and Appendix Figure 1B, available online).

Sensitivity analyses revealed no significant effect of removing quality and single fruit sale variables or of reducing the number of included healthy items on any findings. A full description can be found in Appendix 3 (available online).

**DISCUSSION**

This study developed an adapted CNE tool to be more appropriate for general population health research by including foods representative of national dietary trends and that are targeted in public policies aimed at reducing energy and sugar intakes in England. The adapted CNE tool showed acceptable inter-rater reliability. Large stores consistently offered more healthful environments than small stores. The tool was sensitive to longitudinal changes in the environmental variables that contribute to store healthfulness.

This study provides novel insights into the healthfulness of supermarket in-store environments. The findings are consistent with the monthly fluctuations in food marketing practices expected throughout January–May that accompany seasonal events. From January to February, there was a greater increase in promotions on healthy items than on less healthy items, likely reflecting the end of promotions on Christmas items and a new focus on supporting New Year health resolutions. In March, the promotions on less healthy items increased, and promotions on healthy items decreased, likely capturing the introduction of highly promoted Mother’s Day and Easter products. Simultaneously, the availability of healthier alternatives of less healthy items decreased, perhaps because healthier alternatives were removed to increase capacity for Easter products. After Easter, variety and promotion of healthy items increased in May because new British spring/summer fruits and vegetables became available and were promoted.

Results consistently showed that large stores offered more healthful environments than small stores, a feature also reflected in majority of the individual variables, suggesting that the variety, quality, promotions, shelf placement, store placement, nutrition information, healthier alternatives, and single fruit sale are better for healthy products in larger stores. This is consistent with previous research conducted in Hampshire, UK, which assessed the healthfulness of 601 retail food stores using the original CNE tool, as well as with other research from high-income countries. A potential explanation for the more healthful environment in large stores is the greater space to stock a wider variety of healthy products and the ability to promote these. Small stores by comparison have limited scope to increase their range beyond core products or to alter where products are placed and promoted. These discrepancies by store type may be contributing to the dietary inequalities, although there was no evidence of an association between store healthfulness and neighborhood deprivation after adjustment for store type. Similarly, research in Scotland, England, and Australia has shown little variation in price and availability by level of neighborhood deprivation, whereas others have reported poorer fruit and vegetable quality and greater promotion of less healthy foods in stores in less affluent areas. Unexpectedly, in small stores, the difference in food prices between healthy and less healthy food products was minimal, showing that they were more supportive of affordable healthy foods, whereas in large stores, cheaper prices favored less healthy items. Collectively, these findings suggest that targets for future intervention for researchers and
policymakers may differ according to store format, whereby strategies in large stores need to correct the price differential between healthy and less healthy products and small stores could focus on increasing their varieties of fruits and vegetables in an effort to address inequalities.

Numerous tools have been developed to assess supermarket in-store environments; however, few have undergone reliability or validity testing, and most tools focus on measuring 2 specific factors: availability and price. The adapted CNE tool is a comprehensive tool with moderate-to-excellent κ for almost all variables, except for in-store placement and quality. The former could be improved with more precise definitions of in-store locations, such as those used in the GroPromo tool. Fruit and vegetable quality showed the lowest agreement, consistent with findings from previous research. However, removing quality and single fruit sale variables (which showed minimal variation between stores) from the tool had no impact on composite healthfulness scores. Future assessment using photographs or a 2-point scale of acceptable/unacceptable may provide a more robust measure of quality.

The tool detected changes in the healthfulness of certain in-store variables over time and differentiated between different store types in meaningful ways. Ways to further simplify the tool without materially changing the findings were also identified. For example, removing 4 healthy items to match the number of healthy and less healthy items included in the tool did not materially alter the findings and will reduce the burden of future assessments. By assessing products targeted by the government sugar and calorie reduction programs, the tool could be useful in providing information at a store level on the healthfulness of marketing strategies or the change over time. This adapted tool could be used to conduct a wider investigation of the specific differences in healthfulness between stores and over time. Furthermore, it could be used to monitor the effectiveness of policies intended to improve the healthfulness of in-store environments.

This study expanded on research by Black et al. by adapting their existing tool to include greater emphasis on products that are the focus of current policies to improve the healthfulness of the food supply and by assessing changes in in-store variables over time. However, not all stores had data collected every month, and the longitudinal analyses only included stores with ≥2 observations within the study period, considerably reducing the sample size and the ability to detect differences in store healthfulness that may exist across deprivation levels. Considering known regional differences in food environments and that London has on average a younger, more educated population with greater earnings than the rest of the UK, which may influence supermarket design and strategy and possibly impact store healthfulness, the generalizability of the results is limited.

The complexity of supermarket environments makes it difficult to measure all aspects of the environment that may influence purchasing decisions, but changes in the tool variables may be offset by other in-store changes. The tool is intended to give an overall store healthfulness rating; however, there is a risk of retailers focusing on specific components of the score while continuing less healthful practices in areas missed by the tool. Future research could link changes in healthfulness scores to changes in the nutritional quality of food purchases to assess the extent to which this occurs.

Finally, given the natural seasonal variations of price and availability on fresh produce, only including fruits and vegetables as healthy items may have limited the ability to detect useful information on pricing and the promotions of other nutrient-dense ambient products. Further research could examine other factors, including a wider range of healthy items. Despite these limitations, this study provides proof of concept that the adapted CNE tool can be used to assess store healthfulness.

CONCLUSIONS

This study provides proof of concept that the adapted CNE tool is able to detect changes over time in specific in-store environmental features that can influence food purchasing decisions. It shows large stores to be more healthful than small stores, particularly given the wider variety of healthy items available and the greater ability to promote these. This knowledge could be useful in identifying areas for change to enhance the healthfulness of smaller store environments.

ACKNOWLEDGMENTS

The authors would like to acknowledge the funding received from Guy’s and St Thomas’ Charity (Grant EIC181003). The authors also like to acknowledge the support from Public Health England and The Food Foundation in recruiting data collectors. The authors would like to thank the data collectors themselves for conducting the store surveys and the
participating supermarkets and store managers for allowing the data collectors to take the necessary measurements for this research.

The study sponsor had no role in the study design; collection, analysis, and interpretation of data; writing of the report; or the decision to submit the report for publication. The views expressed in this publication are those of the author(s) and not necessarily those of the National Health Service, National Institute of Health Research, or the United Kingdom (UK) Department of Health and Social Care.

GH, CP, and SJ are funded by the NIHR Applied Research Collaborations Oxford. CV and GN are supported by funding from the UK Medical Research Council (MC_UU_12011/4) and the National Institute of Health Research Public Health Research Programme (Grant Funding, 17/44/46). SJ is funded by the NIHR Oxford Biomedical Research Centre (BRC) Obesity, Diet and Lifestyle Theme.

CP, CV, and SJ participated in the design of the study. GH and CP adapted the audit tool with input from GN, CV, and SJ. GH and CP coordinated the data collection, healthfulness score development, and analyses and wrote the first draft of the manuscript. GN and CV contributed to the development of the healthfulness score and performed some of the statistical analyses. SJ, GN, and CV provided critical revisions to the manuscript. All authors read and approved the manuscript.

CV has a non-financial research relationship with a UK food retail company and maintains independence in all evaluation activities. However, this article is not related to this relationship. All other authors declare that they have no competing interests. No financial disclosures were reported by the authors of this paper.

SUPPLEMENTAL MATERIAL

Supplemental materials associated with this article can be found in the online version at https://doi.org/10.1016/j.amepre.2021.04.012.

REFERENCES


