Effect of different front-of-package food labels on identification of unhealthy products and intention to purchase the products—A randomised controlled trial in South Africa

Makoma Bopape\textsuperscript{a,b,*,} Jeroen De Man\textsuperscript{c}, Lindsey Smith Taillie\textsuperscript{d}, Shu Wen Ng\textsuperscript{d}, Nandita Murukutla\textsuperscript{e}, Rina Swart\textsuperscript{f}

\textsuperscript{a} Department of Human Nutrition and Dietetics, Faculty of Health Sciences, University of Limpopo, South Africa
\textsuperscript{b} Faculty of Community and Health Sciences, School of Public Health, University of the Western Cape, South Africa
\textsuperscript{c} Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium
\textsuperscript{d} Carolina Population Center and Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, USA
\textsuperscript{e} Vital Strategies, New York City, New York, USA
\textsuperscript{f} Department of Dietetics and Nutrition, Faculty of Community and Health Sciences, University of the Western Cape, South Africa

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Front-of-package food labels
Unhealthy products
Purchasing intention
Obesity
Noncommunicable diseases

\textbf{ABSTRACT}

This study aimed to evaluate the effect of different labels on participants identifying products high in nutrients of concern; identifying unhealthy products, and intention to purchase unhealthy products. This blinded randomised controlled trial included a representative sample of South African households ($n = 1951$). Per household we selected a member primarily responsible for food purchases. Participants were randomised into the Warning Label (WL), Guideline Dietary Amounts (GDA) or Multiple Traffic Light (MTL) arms. Each participant answered questions in a no label condition (control) followed by same questions in the label condition (experiment). Complete data were collected and analysed for 1948 participants ($WL = 33.7\%$, $GDA = 32.1\%$ and $MTL = 34.2\%$). The probability of correctly identifying products high in nutrients of concern and identifying products as being unhealthy was higher with the WL compared to the GDA or MTL for most items. There was no difference in performance between the GDA and the MTL when considering all items together. A higher percentage of participants reported a lower intention to purchase an unhealthy product after exposure to the WL compared to MTL for 5 out of 6 products; 2 out of 6 products for the WL compared to GDA and 2 out of 6 products for GDA compared to MTL. Compared to the control condition, exposure to each of the labels resulted in better identification of nutrients of concern, unhealthy products and a lower intention to purchase when considering all specific outcome items together. The WL showed a higher potential to enable South African consumers to identify products high in nutrients of concern, identify unhealthy products and discourage purchasing of unhealthy products.

\textbf{1. Background}

The prevalence of obesity and noncommunicable diseases (NCDs) is high in South Africa and continues to increase substantially (\textit{Statistics South Africa, 2017}; \textit{WHO, 2018}). Among the South African population, more than two-thirds of women and approximately one-third of men are overweight or obese (\textit{Statistics South Africa, 2017}). Obesity and NCDs are leading causes of morbidity and mortality and have recently been associated with severe complications of infectious diseases such as COVID-19 (\textit{The Lancet, 2020}; \textit{World Obesity Federation, 2021}). Unhealthy diets that are high in energy are among the main causes of obesity and NCDs (\textit{WHO/FAO, 2003}) and effective policies are needed to improve populations food intake and to address these conditions.

In South Africa and across the globe, consumers are continually being exposed to ultra-processed foods (\textit{Baker et al., 2020}; \textit{Moodie et al., 2013}; \textit{Puoane et al., 2012}; \textit{Reardon, Timmer, Barrett, & Berdegué, 2003}) resulting in diets of poorer quality (\textit{Igumbor et al., 2012}; \textit{Imamura et al., 2015}; \textit{Koiwai et al., 2019}; \textit{Monteiro, Moubarak, Cannon, Ng, & Popkin, 2015}).
These unhealthy products are typically high in energy, sugar, saturated fats and salt (Monteiro et al., 2021) which are nutrients linked to the development of obesity and diet-related NCDs (WHO/FAO, 2003). Policies need to be put in place to educate consumers about the nutritional composition and negative health consequences of these unhealthy foods.

The World Health Organisation (WHO) recognises front-of-package labelling (FOPL) as a means to provide accessible, simple and easily understandable nutrition information (WHO Regional Office for Europe, 2020). The principal aim of FOPL is to provide interpretive, attention grabbing and easily understandable nutrition information presented at the point of decision-making to assist all consumers to make informed food purchases and healthier dietary choices (WHO Regional Office for Europe, 2020). There is consensus that the presence of FOPL assists consumers to make a distinction between healthy and unhealthy foods.

In terms of the definition of ‘unhealthy foods’ within the South African policy context, the existing food labelling regulation (R146) only limits itself to the definition of health claims (National Department of Health, 2010). However, draft R429 (Guideline 14), currently in review, includes guidelines on the criteria for the commercial marketing of foods and non-alcoholic beverages to children. The proposed guideline mentions the aim of the Department of Health which is to restrict marketing of unhealthy foods and drinks to children. The document defines unhealthy foods as products high in fat, saturated fats, trans-fatty acids, free sugars, and sodium (salt) (National Department of Health, 2014). The term unhealthy foods is therefore a familiar concept within the regulatory and policy frameworks in the country. Additionally several studies have been conducted in South Africa where the term ‘unhealthy foods’ was used, so this definition of ‘unhealthy foods’ reflects the current state of understanding in this context (Mchiza, Temple, Steyn, Abrahams, & Clayford, 2013; Temple, Steyn, Myburgh, & Nel, 2006; Yamoah, De Man, Onagbiye, & Mchiza, 2021).

The existing FOPL formats, however, differ in their level of complexity and some may be more effective in conveying the healthiness of products based on their design and the level of information included (EUFIC, 2017; European Commission et al., 2020; Ikonen, Sotgiu, Aydinli, & Verlegh, 2020). The food industry in South Africa currently applies a voluntary GDA that consumers in other studies report to be challenging and confusing due to information overload and the technical terms used (Deliza, de Alcantara, Pereira, & Ares, 2020; Egnell, Talati, Hercberg, Pettigrew, & Julia, 2018). This underlines the need for a simpler FOPL format that can easily convey nutrient information within the South African context (Koen, Blaauw, & Wentzel-Viljoen, 2016). South Africa has not implemented an FOPL system and is in the process of updating the current Regulation 146 (R146) to include FOPL (National Department of Health, 2010, 2014).

The selection of the FOPL should consider the country’s unique context including the educational and income status of the population for it to be effective and equitable. FOPLs vary by format and design and can be classified as either reductive or interpretive (EUFIC, 2017; Ikonen et al., 2020; Kanter, Vanderlee, & Vandevijvere, 2018; Kelly & Jewell, 2018). Reductive FOPL systems such as the Guideline Dietary Amounts (GDA) (Fig. 1) inform consumers by highlighting nutrients associated with NCDs on the front of pack without providing any conclusion about the healthiness of the product (Food and Drink Federation, 2013; Ikonen et al., 2020).

The GDA appears on all food products irrespective of their nutritional quality (Hodgkins et al., 2012) and requires consumers to make judgements about the healthiness of the products. Due to the numerical interpretations required to understand the information on the label, reductive FOPLs tend to be the worst performing in informing consumers about the healthiness of food products (Deliza et al., 2020; Vargas-Meza, Jaüregui, Contreras-Manzano, Nieto, & Barquera, 2019). The GDA is currently voluntarily applied on the majority of packaged foods in South Africa (Igumbor et al., 2012) and may be familiar to South African consumers.

Interpretive nutrient specific systems such as the Multiple Traffic Light (MTL) and Warning Labels (Fig. 1) evaluate the nutritional quality of products by using interpretive aids such as colour, icons and shapes

![Guideline Daily Amounts](image1)

![Multiple Traffic Light (UK)](image2)

![Multiple Traffic Light (Ecuador)](image3)

![Warning Label (South Africa tested in this study)](image4)

Fig. 1. Examples of front-of-package labels.
excessive (Grummon et al., 2019; Kelly et al., 2019) and that consumers find the MTL attractive due to its colour combinations. However, MTL may be less helpful in assisting consumers evaluate the healthfulness of a product in cases where a product carries a different color for each nutrient and thus providing conflicting messages (Khandpur et al., 2018; Gorski Findling et al., 2018; Machín, Aşıklı, Contreras-Manzano, & Barquera, 2019). A recent study eliminated MTL as a potential FOPL for testing in South Africa following consumers’ suggestions to implement black and white warning labels than colored FOPL (Todd, Guetterman, Volschenk, Kidd, & Joubert, 2022). Similar to the GDA, the MTL appears on foods irrespective of their overall nutritional value (Hodgkins et al., 2012) making it challenging to judge if a product is healthy or not especially in instances where each nutrient is allocated a different color. However, MTL has consistently outperformed the GDA in improving consumers understanding of the nutritional quality of products (Arrúa, Machín, et al., 2017; Egnell et al., 2018; Gorski Findling et al., 2018; Khandpur et al., 2018).

Warning Labels (WL) (Fig. 1) are another type of FOPL that use colour, pictorial images and texts such as ‘high in’ or ‘excessive’ to interpret the products nutritional information (Chile Ministry of Health, 2015; WHO Regional Office for Europe, 2020). WLs have been found to outperform other labelling formats in assisting consumers identify unhealthy products (Talati, Egnell, Hercberg, Julia, & Pettigrew, 2019) and in reducing consumers intention to purchase unhealthy products compared to other labelling formats (Khandpur et al., 2018; Talati et al., 2018). Results of a qualitative study in South Africa reveal that consumers found the WL simple and easy to understand (Bopape et al., 2021). The WL scheme aims to, at a glance, highlight and discourage consumptions. However, WLs may be less helpful in assisting consumers evaluate the nutritional quality of products (Khandpur et al., 2018; Grolnick et al., 2018). The WL design was based on the results of a qualitative study conducted in 2019 among South African consumers of varying socioeconomic, demographic and educational backgrounds (Bopape et al., 2021).

To our knowledge, no studies have been conducted in the African region, including South Africa, to compare and identify the FOPL that best enhances nutrient understanding among the countries’ general population. This randomised trial aims to fill this gap by comparing the performance of three different types of FOPL in samples representative of South African households. Specifically, we will measure if these labels assist South African consumers to i) identify products high in nutrients of concern (i.e. saturated fats, sugar and salt); ii) identify unhealthy products and; iii) reduce their intention to purchase unhealthy food. This study will focus on the existing FOPL systems (GDA and MTL) and the proposed WL designed for South Africa, using their respective nutrient profile models (NPMs) and label design.

The findings of this study will provide evidence for public policies at national, regional and international levels that aim to inform and assist populations in making healthier dietary choices.

2. Methods

This study was a three-armed randomised controlled trial (RCT) with both a within and between subject factor. The within-subject effect corresponded to the difference between a no-FOPL vs. FOPL product. The between-subject effect corresponded to the difference between the three FOPL conditions. The reporting of the methodology was based on the CONSORT guidelines for reporting parallel group randomised trials (Schulz, Altman, & Moher, 2010). This trial was pre-registered with ‘As Predicted’: 45567

2.1. Sampling strategy and sample size

A stratified multistage random sampling strategy was used to obtain a representative sample of the South African population. Primary sampling units were pre-determined Enumerator Areas (EAs) which were proportionally stratified for: 1) geographical area (metro urban/metro traditional/non-metro urban/non-metro traditional), 2) socioeconomic status (low, middle and upper income categories), 3) province and 4) population groups. Socioeconomic status was based on the Neighbourhood Lifestyle Index® (NLI®) developed by an independent statistician and used by a geographical information service (GEOTERRAIMAGE (GTI)) (https://geoterrimage.com/neighbourhood-lifestyle-index/) that provided maps for the survey. The NLI® is a system modelled from population dwelling unit information classifying neighbourhoods according to their income and various lifestyle characteristics from 1 (lowest income/poorest community) to 10 (highest income/most affluent community). For this survey, NLI® was categorised into three wealth status groups (low, middle, and upper income) groups. Per selected EA, 15 households were selected as secondary sampling units through interval sampling.

The sample size of 1526 households was calculated a priori at a power of 90%, estimated effect size of 0.136 (Duboc, 2015) and 95% confidence level. Although the latter study assessed acceptability, we could expect such an effect size in our study which is deemed relevant at population level. The sample was overestimated to 2500 to account for possible non-responses and hard to reach residential areas.

2.2. Participants’ recruitment

Data were collected at the participants’ households. Within each household the person primarily responsible for food purchases was selected if aged ≥18 years and after consent was obtained. If this person was not present, then somebody who was co-responsible for food purchases was selected. If the latter was not present, then someone who occasionally assisted with food preparation was selected.

Participants’ recruitment and data collection were conducted by a research agency with extensive experience in data collection. Fieldworkers were trained on how to select households, recruit participants and how to administer the questionnaire. Data collection took place between 29 November–12 December 2019 and 6–31 January 2020. Ethical approval was granted by both the Biomedical Research Ethics Committee of the University of the Western Cape and the International
Research Board at the University of North Carolina (Chapel Hill). Areas in the upper income category were inaccessible (many resided in gated communities) and therefore excluded from the study. The exclusion of this small proportion of households was deemed acceptable as FOPL is deemed more relevant to lower socioeconomic groups that are harder to reach through education channels and other media (see Fig. 2 for a diagrammatic presentation of the number of participants that enrolled in the study).

2.3. Procedure

Participants were randomised to one of three study conditions: products containing a WL, products containing a GDA, or products containing MTL. Each fieldworker was assigned one label type on the day of data collection (either the WL, GDA or MTL) which were rotated daily and participants’ label exposure depended on the label that the fieldworker applied on the day. Randomisation depended on the number of recruitment possible on the day of data collection. This procedure resulted in 33.7% (n = 656) being exposed to the WL, 32.1% (n = 626) to the GDA and 34.2% (n = 669) to the MTL (Fig. 2). Randomisation to the label type took place before participants were enrolled into the study and all the participants were blinded to the label they were allocated to. Blinding fieldworkers was not possible due to the data collection procedures followed.

In each experimental arm participants were first shown images, one at a time, of products without FOPL (control phase) followed by images of the same products bearing one of the three labels participants were allocated to (experimental phase) (Fig. 3). Each participant was exposed to both the control and experimental phase with the aim to determine the within and between subject effect. In both phases, participants were asked questions after they viewed the products. The questionnaire used in this study was adapted from a similar RCT study conducted in Brazil (Khandpur et al., 2018) (see Additional File 1). The adapted questionnaire was piloted on 10 post-graduate students from different language and socio-economic backgrounds for validity, logical flow, and clarity of questions. The questionnaire was further tested in two EAs on 21 individuals from low sociodemographic background. The pilot study resulted in the nutrition knowledge questionnaire (not part of this paper) being shortened as the questionnaire took too long to administer, unclear questions were rephrased and repeating questions were deleted (also not part of this paper). In addition to the primary outcomes, data on food consumption patterns, demographics and socio-economic status were collected. Data was collected using smartphones.

2.4. Stimuli

Products: Products used in this study included fictitious images of crisps, fruit juice and soda (single products) and two packets of sweet biscuits, cereal and yoghurt with different brand names (paired products) (see Additional File 2). The researchers developed four sets of fictitious products containing all nine products: one set without FOPL (control condition), and three with one of each FOPL: WL, GDA or MTL (see Additional File 2 for all products). The same products were used in both the control and the experimental phases. The selection of product categories was informed by top sales in South Africa in 2018 according to Euromonitor and were meant to represent a mix of products often perceived as unhealthy (crisps, sweet biscuits and soda) and products whose healthfulness is more ambiguous (100% fruit juice, cereals and yoghurt). The use of fictitious products was chosen to minimise bias due to participants’ preconceived knowledge, product familiarity and brand preferences. All participants saw the same sets of products and the only variation was on the labels that were applied. Each product pair contained one product with lesser amounts of nutrients of concern.

The nutritional information on the products was based on similar commercial products available on the South African markets. The information was based on per 100 g/ml and these nutritional profiles were the same across the FOPL. The labels were allocated according to this information. The labels contained information about saturated fats, sugar, salt and artificial sweetener in the case of yoghurt in the WL arm. The nutritional information and interpretive information for each product is presented in Additional File 3). All the labels were placed on the top right corner of each food package.

Labels: The FOPL tested in this study were the GDA (Food and Drink Federation, 2013), MTL (FSA, 2016) and the WL designed for use in

![Fig. 2. Diagrammatic presentation of study enrolment and randomisation.](image-url)
South Africa (Bopape et al., 2021) (Fig. 1). Real world FOPLs with their respective NPMs were used as a way to assess policy options in discussion. However, the lack of disentanglement of the NPM and the label design may create challenges in determining whether the outcomes are due to the underlying NPM or the label design itself. Energy (kJ), sugar (g), fat (g), saturates (g) and sodium (mg) content were presented per 100 g/ml for the GDA and MTL for each product. The WL only appeared on products that contained excessive amounts of nutrients of concern. The % Reference Intakes (RI) was based on the estimated requirements for a 70 kg adult (2000 kcal) (Food and Drink Federation, 2014). Nutritional interpretation for the MTL included the use of colour codes reflecting low (green), medium (amber) and high (red) nutrient content (FSA, 2016). The nutrient criteria for the WL were based on the proposed South African Nutrient Profile Model (Frank et al., 2021) and warnings were shown using a triangle sign.

2.5. Ethics

Ethical principles were applied in the execution of this study in accordance with the Declaration of Helsinki. Informed written consent was obtained from all participants before data collection commenced. Ethical approval was granted by both the Biomedical Research Ethics Committee of the University of the Western Cape (BM 18/9/13).

2.6. Outcome measures

For single products (crisps, juice and soda) assessment, the primary outcomes were whether the participant correctly identified products that were high in salt, sugar and saturated fat (yes/no/don’t know) and correctly identified products as unhealthy (healthy/unhealthy). All products used in this study were unhealthy. A product was considered high in nutrients of concern or unhealthy if it either contained a WL or two colours on the MTL were either amber or red. For the paired-product (biscuits, cereal and yoghurt) assessment, the primary outcomes were whether the participant correctly identified the product that was higher in salt, sugar, or saturated fat; and whether the participant correctly identified the unhealthier product. For paired products, a product was considered higher in nutrients of concern or unhealthier if it had more WLs or either more amber or red than the green colours. In a case where the MTL had similar colour patterns, the unhealthiness was determined by the differences in nutrient amounts.

We examined change in intentions to purchase unhealthy products with the question: “How likely are you to buy this product for yourself or your family?” The responses were based on a four-point Likert scale with the following options: “I would definitely not buy it”; “I am unlikely to buy it”; “I will consider buying it”; “I will definitely buy it”. All responses were collapsed into binary responses – 1 = Yes, 0 = No and “Don’t know” responses were recoded as = 0.

2.7. Statistical analysis

To compare the effect between labels, a modified Poisson regression model was used with the follow-up (i.e., the product shown with one of the three labels), outcome measures as the dependent variable and the different labels as independent variables. To account for a potential regression to the mean effect, “analysis of covariance” was conducted with the baseline (control) value included as a covariate. As recommended for real-world trials (Kahan, Jairath, Doré, & Morris, 2014), we adjusted for potential confounders including: age, sex, level of education, socio-economic status, being the main responsible buyer, having children and metropolitan residence. Results were presented as relative risk (RR) estimates comparing two labels. RR > 1 implied a higher percentage of participants exposed to label X identified products high in nutrients of concern or unhealthy products correctly compared to participants exposed to label Y.

To measure the absolute effect for each label separately (i.e., within-subject differences), a paired comparison was used to calculate the difference between the proportion of right answers at baseline (i.e., the product was shown without a label) and follow-up (i.e., the product was shown with one of the three labels). Standard errors were calculated taking into account the paired design. In addition, the proportion of change from ‘yes’ to ‘no’ or vice versa was calculated between both data collection points.

For both between and within effect calculations, survey design weights and clustering at the level of the EAs were taken into account. Cases with missing data were rare (N = 3) and deleted listwise. R software was used, with the packages “geepack”(Halekoh, 2006) and “survey” (Lumley, Maintainer, & Lumley, 2021). The hypotheses were specified before the data were collected, the analytic plan was pre-specified and any data-driven analyses are clearly identified and discussed appropriately.

3. Results

Table 1 presents the socioeconomic profile of the participants according to the FOPL type.

Fig. 3. Example of images used as stimuli during data collection.
3.1. Identification of products high in nutrients of concern

Fig. 4 presents the differences (in relative risks) between the different FOPLs. The probability of correctly identifying products high in nutrients of concern was almost twice as high for certain products when exposed to the WL than to either the GDA or the MTL (black squares in Fig. 4). For example, the probability of correctly identifying that biscuits were high in fat was 1.85 (Confidence Interval (CI) = 1.60–2.14) times higher after exposure to the WL vs. to the GDA and the probability of correctly identifying that juice was high in sugar was 1.67 (CI = 1.47–1.89) times higher after exposure to the WL vs. to the MTL (Fig. 4). The relative risks comparing the WL and the MTL showed a higher variability (Fig. 4) than the ones comparing the WL and the GDA with a minimum of 1.03 (CI = 0.96–1.12) for yoghurt high in sugar up to a maximum of 2.64 (CI = 2.15–3.25) for biscuits high in fat. Relative risks comparing the GDA and the MTL (Fig. 4) did not indicate exposure to either label as more advantageous to a correct identification over the whole range of products.

3.2. Identification of unhealthy products

Similarly, a higher percentage of participants exposed to the WL correctly identified unhealthy products compared to the GDA and MTL groups (red dots in Fig. 4). When comparing exposure to the WL versus the GDA, the probability of correctly identifying unhealthy products was higher for all products and when comparing exposure to the WL versus the MTL, the probability was higher for all products except for biscuits (Fig. 4). For example, the probability of correctly identifying that yoghurt was unhealthy was 1.64 (CI = 1.4–1.91) times higher with the WL vs the GDA and was 1.97 (CI = 1.53–2.54) times higher with the WL vs the MTL for juice.

Table 1
Participants’ sociodemographic information by FOPL type (n = 1948).

<table>
<thead>
<tr>
<th>Age Mean (SD)</th>
<th>WL 655</th>
<th>GDA 626</th>
<th>MTL 667</th>
<th>TOTAL n = 1948</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>37 (14)</td>
<td>38 (15)</td>
<td>36 (14)</td>
<td>37 (14)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Socio-economic status</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>430</td>
<td>66</td>
</tr>
<tr>
<td>Middle</td>
<td>181</td>
<td>27</td>
</tr>
<tr>
<td>High</td>
<td>44</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Urban residence</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>595</td>
<td>91</td>
</tr>
<tr>
<td>No</td>
<td>60</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>228</td>
<td>35</td>
</tr>
<tr>
<td>Females</td>
<td>427</td>
<td>65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Educational level</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary (&lt;Grade 7)</td>
<td>38</td>
<td>6</td>
</tr>
<tr>
<td>Secondary (Grades 7–11)</td>
<td>257</td>
<td>39</td>
</tr>
<tr>
<td>Grade 12</td>
<td>202</td>
<td>31</td>
</tr>
<tr>
<td>Tertiary</td>
<td>158</td>
<td>24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary grocery buyer</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>51</td>
<td>8</td>
</tr>
<tr>
<td>Yes</td>
<td>376</td>
<td>57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Share responsibility</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>228</td>
<td>35</td>
</tr>
<tr>
<td>Yes</td>
<td>427</td>
<td>65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Children &lt; 18yrs present</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>143</td>
<td>22</td>
</tr>
<tr>
<td>Yes</td>
<td>512</td>
<td>78</td>
</tr>
</tbody>
</table>

WL = Warning Label; GDA = Guideline Dietary Amounts; MTL = Multiple Traffic Light.
3.3. Intention to purchase

Fig. 5 shows that the probability of expressing the intention to purchase unhealthy products was lower for most products bearing the WL than either the GDA or MTL (Fig. 5). For example, the probability of expressing the intention to purchase chips and soda for the group exposed to the WL was 0.75 (CI = 0.67–0.85) and 0.81 (CI = 0.71–0.92) times the probability for the group exposed to the GDA. The MTL performed worst with a higher RR than the GDA in more than 50% of the products.

3.3.1. Absolute change per label for the identification of nutrients of concern

We found a substantial increase in the proportion of participants identifying products as high in nutrients of concern after exposure to the WL for all products, except for biscuits high in sugar (See Additional File 5). The largest change in correctly identifying nutrients of concern after exposure to the WL was recorded for chips high in fat (37%), for biscuits high in fat (23%) and for juice high in sugar (23%). For the other labels, we also noticed an increase in several, but not all of the outcomes and point estimates were lower compared to the WL.

When comparing the proportion of participants that changed from incorrect to correct identification or vice versa, there was a relative difference between the labels. When comparing the GDA and MTL labels to the WL, a higher proportion of participants changed from correct identification to incorrect identification for the GDA and MLT labels. For example, for chips high in fat for the WL, 41.1% changed from incorrect to correct identification, while 4.5% changed from correct to incorrect identification. Subtracting both figures (41.1 – 4.5%) gives the actual difference pre-post: 36.6%. For the GDA, however, the proportion that changed from incorrect to correct was 29.0% and the proportion that changed from correct to incorrect was 14.7 (See Additional File 4).

3.4. Absolute change per label for the identification of unhealthy

Identification of unhealthy products improved for 4 out of 6 products after exposure to the WL; no significant difference was shown for chips and biscuits (See Additional File 4). The proportion of participants who correctly identified unhealthy products post-exposure to the WL improved by 24% for juice, 23% for cereal, 19% for yoghurt and by 7% for soda. Exposure to the GDA and MTL only improved identification for juice, cereal and yoghurt and estimates were lower than for the WL. Exposure to the GDA and MTL resulted in a substantial decrease in correct identification of chips being unhealthy. Similarly to nutrients of concern, we saw a relatively higher trend to change from correct to incorrect identification for the GDA and the MTL compared to the WL (See Additional File 4).

3.5. Absolute change per label for intention to purchase

Reported intention to purchase unhealthy products was significantly reduced post-exposure to the WL for all products. The GDA and the MTL showed a reduction for most, but not all products and the differences pre-and post-exposure were smaller (See Additional File 5).

4. Discussion

The findings of this study indicate that, compared to the GDA and the MTL, the WL mostly performed well in all three outcomes: assisting consumers identify products high in nutrients of concern, identifying unhealthy products and reducing intention to purchase unhealthy products. Overall, the GDA and the MTL also did facilitate some changes...
in the three outcomes, but the effects were weaker and occurred in fewer of the studied products.

These findings for South Africa are consistent with other experimental studies. For example, a study in Brazil reported that more participants were able to identify products with excess nutrients in the presence of the WL than with the MTL (Khandpur et al., 2018). Moreover, other studies have reported challenges with consumers use and understanding of GDA and MTL (De la Cruz-Gongora et al., 2017; Vargas-Meza, Jairegui, et al., 2019). A study in Uruguay reported that the GDA and the MTL performed the same in their evaluation of the healthfulness of products (Arrúa, Machín, et al., 2017).

Earlier studies have found that while the MTL performed better in assisting consumers identify healthier products (Taillie et al., 2020; Vargas-Meza, Jairegui, et al., 2019), consumers still found the MTL challenging to interpret when the label contained two or three different colours (Grunert, Wills, & Fernández-Celemin, 2010; Kees, Royne, & Cho, 2014; Machín et al., 2018). Indeed, we found that for products that contained all three traffic light colours on the labels such as chips and biscuits used in this study, the MTL performed badly in assisting consumers to identify these products as being high in nutrients of concern. Chips, for example, contained an amber colour for sodium, which could have been difficult for the participants to interpret. The WL, due to its binary nature, unlike the GDA or MTL, only highlighted that chips were high in salt without mentioning other nutrients. Warning Labels, due to their single attribute nature, present concise and easy to interpret information and may therefore be more effective in informing consumers. It is however not clear if the difference in outcomes were due to the labels or their NPMs or how much is due to which.

4.1. Identifying unhealthy products

Our findings that the WL assisted more participants to identify unhealthy products are also consistent with past experiments in other countries (Arrúa, Curutchet, et al., 2017; Arrúa, Machín, et al., 2017; Khandpur et al., 2018; Neal et al., 2017; Newman, Burton, Andrews, Netmeyer, & Kees, 2018). A recent South African study reported similar findings (Todd et al., 2022). This is likely because the WL, in contrast to the GDA and the MTL, may have simplified and therefore guided consumer identification of unhealthy products (Cecchini & Warnin, 2016; Newman et al., 2018) by only displaying nutrients that are in excess and thus cut through the information noise on the front of the packaging. Additionally, the use of the triangle on the warning label, associated with danger or caution in South Africa (Bopape et al., 2021) could also have led to increased risk perception and an indication of the unhealthiness of the product (Wogalter, Conzola, & Smith-Jackson, 2002). In comparison to the GDA and the MTL, the size of the WL used in this study was bigger, which could have increased its visibility and risk awareness. However a study by Vargas-Meza, Jairegui, et al., 2019 found no difference between the WL and the MTL in assisting consumers identify products with the lowest nutritional quality.

In other experimental studies, the WL decreased healthfulness perceptions of unhealthy products and assisted participants identify unhealthy products (Khandpur et al., 2018; Lima, Ares, & Deliza, 2018) or refrain from choosing unhealthy snacks (Engell et al., 2019). Due to its simplicity, WL are easy to understand and have been proven to be also effective in children (Correa et al., 2019), youth (Hock et al., 2021) and across educational levels (Pereira, 2019). A study exploring two FOPL showed that the presence of the WL enabled children to avoid the unhealthier snack and to choose the healthier option (Arrúa, Machín, et al., 2017). Our results suggest that the use of simple text and familiar icons such as a triangle within the WL would be particularly useful in a setting with low literacy skills such as South Africa. Additionally, the use of icons may enhance equitable access to nutrition information, given the mix of languages spoken in South Africa.

4.2. Intention to purchase

While intention is not necessarily the same as actual purchasing, it could however precede behavioural change (Grummon & Hall, 2020; Taillie et al., 2020). Findings from this study indicate that all three labels had an effect on reducing the reported intention to purchase unhealthy products, but more participants reported a reduced intention after exposure to the WL compared to either the GDA or the MTL. The WL has been reported to reduce intention to purchase unhealthy products due to its potential to improve nutrient content understanding (Grummon et al., 2019; Jairegui et al., 2020; Song et al., 2021) and possibly through eliciting negative emotions towards unhealthy food (Taillie et al., 2020). The presence of texts such as ‘high in’ and ‘warning’ on the WL signals warning for consumers (Grummon et al., 2019; Lehto & Clark, 1991) and that could deter consumers from purchasing or consuming the unhealthy product (Conzola & Wogalter, 2001; Wogalter et al., 2002). The use of pictures or icons such as a teaspoon full of sugar enhances understanding of the unhealthiness of products especially among low literate groups and is suggested to improve adherence to health messages (Houts, Doak, Doak, & Lollowa, 2006). Participants in a previous qualitative study in South Africa, felt that the use of the text ‘warning’ and the inclusion of an exclamation mark on the WL warned them about the danger linked to consumption of the product (Bopape et al., 2021). Consumers in another study reported that health warnings followed by nutrient warnings on food packages evoked fear and reduced their desire to purchase unhealthy foods (Grummon et al., 2019). However in an experimental study by Machín et al., 2018, both the WL and the MTL led to a similar effect of reduced selection of unhealthy products. The findings of this study are consistent with the results of a meta-analysis that reports decreased healthfulness perception and decreased intention to purchase unhealthy products when exposed to the WL (Grummon & Hall, 2020; Song et al., 2021). Countries that experience high obesity and NCD rates and aim to reduce consumption of unhealthy food could therefore benefit from implementation of the WL that easily flag unhealthy products and steer consumers away from such products.

5. Strengths & limitations

Our study was based on a probability sample, randomly collected from and hence representative for the general population of South Africa. Another strength was that data was collected in person and reached lower socio-economic segments of the population unlike online surveys that may be biased towards a more privileged group. The use of an RCT that compared the three FOPL formats, including the no label condition minimised the influence of confounding factors. An added advantage was that the participants interpreted the different FOPLs without any prior explanation of the labels which could have otherwise influenced the intention to purchase and increased the ability to correctly identify unhealthy products. The limitation of this study is that only three FOPLs were tested. However, the three labelling formats represent different labelling categories and the results should provide policymakers with the information required to determine the label that can assist South Africans identify and reduce the purchasing of unhealthy food. Because we used the actual nutrient profile models for the GDA and the MTL which do not require a warning for artificial sweeteners, the yoghurt only contained a warning for artificial sweeteners in the WL arm as compared to the WL, nutrient profile model implemented in other countries, e.g. Mexico (Grunert & Wills, 2007). This could have influenced the ability of consumers to identify yoghurt as unhealthy. This study did not fully differentiate the effects of the type of label (including colour, and label size) and the different nutrient profiles that underpin the labelling system. It is the case that products for which the underlying NPMs are very unaligned, we are likely to see bigger differences in outcomes (compared to products for which the NPMs are more akin). Because this study did not disentangle the NPM and the FOPLs, we are unable to tell whether the
findings were due to the label designs alone or also due to the underlying NPMs and how much is due to which. We recommend future research that could investigate the effectiveness of FOPL when the NPMs are standardised.

The study was experimental and does not represent a real life scenario. Understanding of nutrient content and intention to purchase may not necessarily lead to behaviour change in a real life situation. Studies that determine the effect of WL on actual purchasing and consumption are therefore required.

6. Conclusion

In South Africa, WL performs better than the GDA and the MTL in enabling consumers to identify products high in nutrients of concern, identifying unhealthy products, and reducing their intention to purchase unhealthy products. Requiring mandatory WL for products as defined by a nutrient profile modelling system suitable for South Africa is a feasible and equitable policy that should be considered urgently as the country develops and updates its food labelling regulations.

Ethics

Ethical principles were applied in the execution of this study in accordance with the Declaration of Helsinki. Informed written consent was obtained from all participants before data collection commenced. Ethical approval was granted by both the Biomedical Research Ethics Committee of the University of the Western Cape (BM 18/9/13) and the International Research Board (IRB) at the University of North Carolina (Chapel Hill).

Authors’ contributions

SN, LTS, NM, RS and MB conceived and designed the study. RS and MB executed the study, processed and cleaned the data. SN, LTS, NM, RS and MB critically reviewed the manuscript. All authors read and approved the final manuscript.

Funding

This work was funded by Bloomberg Philanthropies #5108311. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of this manuscript.

Declaration of competing interest

None.

Data availability

Data will be made available on request.

Acknowledgements

We thank the School of Public Health at the University of the Western Cape and the DSI/NRF CoE in Food Security (UID 91490) for administrative support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.appet.2022.106283.

References


Deloia, R., de Alcantara, M., Pereira, R., & Ares, G. (2020). How do different warning signs compare with the guideline daily amount and traffic-light system? Food Quality and Preference, 80, Article 103821.


M. Bopape et al.

Appetite 179 (2022) 106283


**Glossary**

- **CI**: Confidence Interval
- **FOPL**: Front-of-package Labelling
- **GDA**: Guideline Daily Amounts
- **MTL**: Multiple Traffic Light
- **NPM**: Nutrient Profile Model
- **NCDs**: Noncommunicable diseases
- **R146**: Regulation 146
- **RCT**: Randomised Controlled Trial
- **WHO**: World Health Organisation
- **WL**: Warning Label