Understanding people's judgments of the healthiness of food labels:

Two reasoning processes or one?

Seok-Jun Kang

9461 Words

This thesis is submitted in partial fulfilment of the Honours degree of

Bachelor of Psychological Science (Honours)

Table of Contents

List of Figures

List of Tables

Abstract

Various types of food label systems have been created to promote healthy food choices, but the cognitive processes that underlie healthiness assessments are not well understood. Influential dual-process theories have been applied to help understand how people make healthiness assessments, positing a distinction between Type 1 processing, which is intuitive, and generally faster and more error-prone, and Type 2 processing, which is explicit, generally slower and tends to be more accurate. However, the validity of dual-process theories has been challenged. As an alternative, single-process theories state that a range of judgments (such as fast versus slow ones) are based on a common form of assessment. To experimentally test these competing theories, a two-response task was implemented, with fictional food product stimuli that varied in summary Health Star Ratings (HSRs), detailed Nutrition Information Panels (NIPs) and branding logos. Participants first rated the food's healthiness based on their initial impression of the entire label, as quickly as possible. Participants then made a second healthiness rating based on careful examination of the NIP. Results showed that HSRs and logos have a larger effect on the fast first responses, whereas NIPs have a larger effect on the slower second responses. This is consistent with classic dual-process theories. However, when the data were examined using Signed Difference Analysis, there were no ordinal patterns that were forbidden by a single-process model that was based on signal detection framework. Therefore, such formal single-process models offer a viable account of people's healthiness assessments.

Keywords: Nutrition judgment; dual-process theories; single-process theories; two-response task; signed difference analysis

Declaration

This thesis contains no material which has been accepted for the award of any other degree of diploma in any University, and, to the best of my knowledge, this thesis contains no material previously published except where due reference is made. I give permission for the digital version of this thesis to be made available on the web, via the University of Adelaide's digital thesis repository, the Library Search and through web search engines, unless permission has been granted by the School to restrict access for a period of time.

Contribution Statement

In writing this thesis, my supervisor and I collaborated to generate the research questions and design the appropriate methodology. I conducted the literature research, completed the ethics application, and created the fictional food label stimuli, including generating novel product logos and the Nutrition Information Panels used in the experiment. I set up the questionnaire pages in Qualtrics for both the pilot study and main experiment. I was responsible for participant recruitment and managing the online testing. My supervisor wrote R code for data processing and generating graphs, and I coded all statistical analyses in R and generated all other figures. I wrote up all aspects of the thesis.

Acknowledgements

I would like to express deep gratitude to my supervisor Dr. Rachel Stephens for her guidance, encouragement and support throughout the course of my work. Her expertise, advices and faith in me motivated me throughout every stage of this research. I am also grateful to my parents for their continuous encouragement, support and attention.

Understanding people's judgments of the healthiness of food labels:

Two reasoning processes or one?

Nutrition Judgments and Food Labelling

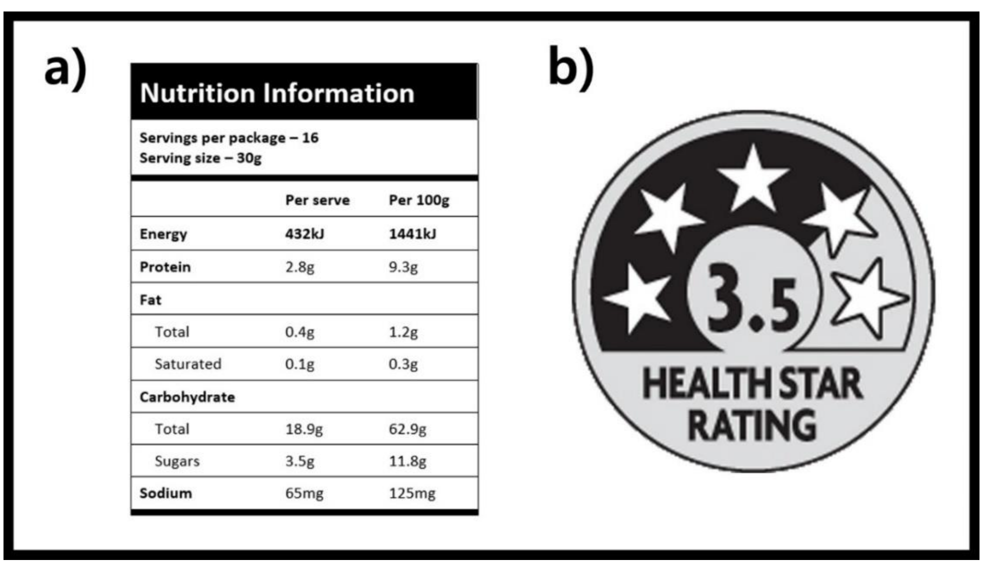
The high prevalence of obesity has become a major health concern for many countries worldwide including Australia (Australian Institute of Health and Welfare, 2019). Leading factors for this poor health condition are the inadequate intake of healthy foods and excess intake of unhealthy foods (Neal et al., 2017). Therefore, a number of policies have been created to promote healthy food choices and better eating habits. However, the perceived nutritional value of food by individuals – how people judge the nutrition quality of foods, and whether those judgments are accurate – remains an important psychological factor of food decisions. This is because, unlike natural foods that are often easily classified as nutritious, such as fruits and vegetables, it could be a difficult challenge to properly comprehend the nutritional value of packaged and/or processed food products.

In many countries including Australia, it is required by law for manufactured foods to carry labels containing safety and nutrition information. For most foods in Australia, a Nutrition Information Panel (NP) must be provided (Australian Government Department of Health, 2013; see Figure 1a). This panel gives detailed and accurate nutritional information to consumers so they can make appropriate food decisions. The biggest shortcoming of Nutrition Information Panels, however, is that it is difficult for people without nutrition guidelines or knowledge to use them properly. For instance, a study by Gnzalez-Vallejo et al. (2016) investigated the direct link between the information present on packaged foods and judgments of nutrition based on the NIP. Results showed that the accuracy of nutrition judgments across participants was low, with a median of only 49% percent of the variance in judgments explained by the nutrients. This highlights that the benefits of the complex NIP label may be quite modest.

To address this issue, there are alternative methods for presenting nutritional information in an easily understandable format. For example, front-of-pack (FOP) interpretive nutrition labels that use graphics and colours to depict nutrient content are likely to be a better option for consumers (Neal et al., 2017). The best-known examples of such interpretive FOP labels include the Australian Health Star Rating (HSR; see Figure 1b) scheme that assigns between 0.5 (least unhealthy) and 5.0 (most healthy) stars to a food, in half star increments. However, research that has examined nutrition label effectiveness is inconclusive and implies that nutrition information does not always promote healthier diets (Aschemann-Witzel et al., 2013). This is because the effectiveness of nutrition label formats is influenced by specific contexts and personal variables such as motivation, nutrition knowledge, and time pressure (Sanjari et al., 2017).

Figure 1.

Examples of Food Nutrition Labels



Note. a) Example of nutrition information panel (NIP). b) Example of Health Star Rating (HSR).

There are many other factors that may also influence customers' nutrition judgments and food choices, such as the type of food, brand name and printed photograph or other graphics on the packaging. Indeed, the visual appearance of the food package and brand has been shown to influence consumers' expectations and judgment (Carrillo et al., 2012; Shepherd et al., 1991). Thus, different consumers may make different nutritional judgments on the same food product, and the same consumer may sometimes judge the nutrition quality of foods accurately and sometimes not (Gnzalez-Vallejo et al., 2016). In order to explain these phenomena, some researchers have applied the concept of dual-process theory, drawn from the reasoning and decision-making literature (Sanjari et al., 2017).

Dual-Process Theories

According to dual-process theories, there are two different types of cognitive processes: Type 1 and Type 2. Type 1 processing is described as fast and heuristic in nature, and can be less accurate. It works unconsciously, favouring a salient option based on the person's preferences and familiarity when making a decision (Bago & De Neys, 2017; Stanovich & Toplak, 2012). Thus, a judgment or choice triggered by Type 1 processing is autonomous and immediate. However, there are times when Type 1 processing faces difficulties, such as when there is no choice that is obvious or familiar to the person amongst the options (Dhar & Gorlin, 2013; Sanjari et al., 2017). This is where Type 2 processing is activated. In contrast to Type 1, Type 2 processing is said to be relatively slow and analytical. It requires a heavy load on working memory and works consciously, and therefore the decision is made in a controlled state. When using Type 2 processing, the person compares contents and aspects of each choice based on one's own motivations and goals, and then makes a decision that is most aligned (Sanjari et al., 2017).

Many factors can influence Type 2 processing including the time available, processing capacity, desired level of accuracy, and fatigue (Dhar & Gorlin, 2013).

There are different views as to how Type 1 and 2 processing interacts. According to classic default-interventionist dual-process theory (e.g., Evans & Stanovich, 2013), Type 1 processing is initialised first. If sufficient time, motivation and cognitive resources are available, Type 2 processing may subsequently intervene, potentially, rectifying an initial incorrect response based on Type 1 processing. In contrast, parallel dual-process models propose that Type 1 and Type 2 processing is activated simultaneously from the start rather than in a serial fashion (Epstein, 1994; Sloman, 1996). However, according to this type of model, Type 2 processing is slower than Type 1 processing, and therefore still has a smaller influence on a fast, pressured response than on a slower response.

Dual-process theories have gained popularity and become a major, influential framework in numerous fields of study, from education and assessment to medical diagnosis and managerial decision making (see Stephens et al., 2020). Some studies have also applied dual-process theories to help understand the relationship between nutrition labels and consumers' response to them (Sanjari et al., 2017). Under Type 1 processing, individuals are more likely to make an appeal-based choice, which means food products are more likely to be chosen based on FOP labels that are easy to understand and are based on a familiar concept (such as traffic lights or scoring stars) (van Harpen et al., 2012; Hersey et al., 2013; Sanjari et al., 2017). Under Type 2 processing, individuals are more likely to make a reason-based choice, which means food products are more likely to be chosen based on label features that contain detailed or numerical information to support the food's healthiness, especially if the individuals' goal is to justify their choice by the nutritional value of the food.

Indeed, careful and comprehensive examination of detailed nutrition information has been linked to Type 2 deep processing. There are many factors within the nutrition information panel that have to be considered to make fully justified choices. For example, the per 100g column and serving size (see Figure 1a) need to be used if comparing nutrients in similar food products, and the recommended healthy value for each nutrient is different (Eat For Health, n.d.). Consumers therefore need to engage with deep information-processing tasks such as extensive searching and recalling knowledge, reading and comprehending numerical and abstract information, and making comparisons (Aschemann-Witzel et al., 2013; Balasubramanian & Cole, 2002; Sanjari et al., 2017).

Criticisms of Dual-Process Theories

Although dual-process theories are an appealing framework to account for nutrition judgments, the validity of these theories has been challenged in the reasoning literature due to several criticisms. For instance, despite the commonly held view that faster Type 1 processing is usually either irrational or intuitive, and slower Type 2 processing is typically necessary for rational decision-making, many studies have suggested that this distinction is not reliable Grayot, 2019; Keren, 2013; (Kruglanski & Gigerenzer, 2011).

One key question is whether people actually typically arrive at a correct final response after making an initial incorrect response. As this is a central pillar of the default-interventionist dual-process theory, many studies have been conducted to evaluate this concept. For instance, De Neys (2006) presented participants with a range of classic reasoning problems, such as evaluating the deductive validity of logical arguments that varied in the validity and believability of the conclusion (e.g., “All dogs have four legs. Puppies are dogs. Therefore, puppies have four legs.”). Correct responses were found to be slower than incorrect (presumably, heuristic) responses. On the surface, this finding seems to be in agreement with the dual-process theories' time course assumption that Type 1 processing is faster and Type 2 processing is slower. However, it is not clear from the results that Type 1 processing was engaged first before Type 2 processing (Bago & De Neys, 2017). In other words, there is no guarantee that the participants generated the incorrect answer first, and then corrected it. When engaging with Type 2 thinking, they might have arrived at the correct response without even contemplating the heuristic incorrect response.

In order to explore this issue, the two-response paradigm has been used in several reasoning experiments (e.g., Bago & De Neys, 2017; Thompson & Johnson, 2014). In this paradigm, participants are initially presented with a reasoning problem (usually with potentially conflicting heuristic or logical/statistical cues) with instruction to respond as quickly as possible. Then, participants are presented with the same reasoning problem again, but this time they are asked to take as much time as they want and to give careful consideration before giving a final response. The assumption here is that the first response should measure more Type 1 processing, while the second response measures more Type 2 processing.

Tellingly, in a two-response study conducted by Thompson et al. (2011), people tended to spend little time to reconsider their initial responses, and rarely changed their response in the second stage. It is important to point out that people's tendency to retain *incorrect* heuristic responses is not an unexpected finding for the classic dual-process point of view; this result may simply indicate failure to engage with the optional Type 2 processing, which leads to incorrect responses. However, it is more difficult for the default interventionist view to explain the situations in which the answer was not changed but both the initial and final responses given were the *correct* logical response. This suggests that the logical response can often be generated quickly and intuitively based on only Type 1 processing. In this case, the standard dual-process theory faces a major challenge (Bago & De Neys, 2017).

Since the standard default interventionist theory is inadequate to explain the above findings, alternative dual-process models could be considered, such as parallel dual-process models (Epstein, 1994; Sloman, 1996). As Type 1 processing and Type 2 processing are engaged in a parallel fashion, correct immediate responses based on Type 2 processing could be generated and hence one might think that the results make theoretical sense under this account. It must be noted that, however, that Type 2 processing is still defined as being slower than Type 1 processing in the parallel model (Epstein, 1994; Sloman, 1996). Additionally, and more importantly, Thompson et al. (2011) also observed correct initial responses even after further limiting Type 2 processing by applying a challenging response deadline and concurrent load task. Therefore, the generation of fast and intuitive logical responses is difficult to explain even with the parallel version of dual-process theory.

Therefore, dual-process theorists have been conducting further research in order to offer a better explanation. For example, more complex hybrid dual-process models of reasoning have been proposed recently (De Neys, 2012; Handley & Trippas, 2015; Pennycook et al., 2015). Combining key features of the serial (default interventionist) and parallel model, these hybrid models assume that parallel intuitive processes generate more than one Type 1 response (e.g., a heuristic and an “intuitive logic” response), which might be followed by more demanding but optional Type 2 processing (Bago & De Neys, 2017).

Single-Process Theories

In light of the difficulties faced by classic dual-process theories, and the blurring of the distinction between Type 1 and Type 2 processes in the more recent hybrid models, alternative single-process theories of reasoning and decision making have been proposed. Single-process theories state that both seemingly deliberative and intuitive judgments are based on a common form of subjective assessment. For instance, Kruglanski and Gigerenzer (2011) proposed that, when we reason, when we judge, or when we make decisions that are fast versus slow or more effortful versus less effortful, there is an underlying common core cognitive process. The accuracy of the decision-maker's response is unrelated to whether the cognition process was deliberate or intuitive. Conscious decisions made based on more information and computation can be less accurate than heuristics made with less effort and neglected information (Kruglanski & Gigerenzer, 2011).

More recently, Stephens and colleagues (e.g., Stephens et al., 2018, 2020) have demonstrated that a quantitative single-process model, based on the signal detection framework, can account for a wide range of reasoning judgments about the validity of classic logical arguments – including fast versus slow judgments, or judgments under working memory load. This model may offer a more parsimonious account than competing dual-process signal detection models. A key implication of the successful single-process model is that judgments made at different speeds (or working memory availability, etc.) differ not in distinct intuitive Type 1 versus deliberate Type 2 assessments of the stimulus, but in shifts in response bias or decision threshold. This signal detection approach to reasoning also offers a useful method for comparing competing single- and dual-process accounts of nutrition judgments, so will be further explained in the next section.

Signal Detection Models of Reasoning or Nutrition Judgments

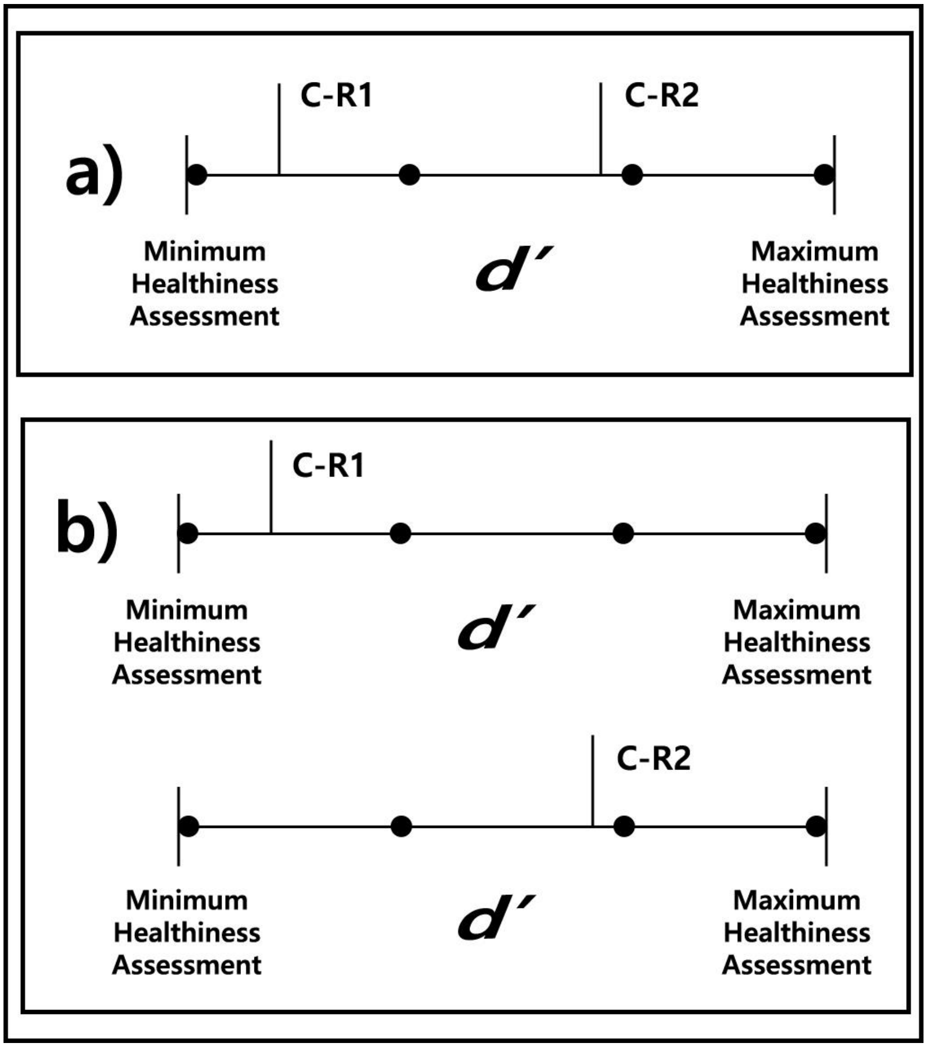
Two important concepts within signal detection theory are: 1) the *discrimination* of stimuli along a subjective strength dimension; and 2) the *decision threshold* (Stephens et al., 2020). In an experiment in which the participants' task is to distinguish "target" and "lure" stimuli (e.g., valid and invalid logical arguments in a reasoning task, or healthy and unhealthy food labels in a nutrition judgments task), the stimuli are assumed to fall along a continuous dimension of subjective strength. Targets and lures are each captured by a different distribution of strength values. The more separated the distributions, the higher the discriminability of the target and lure stimuli. Signal detection models also assume that, when faced with a decision-making situation, a decision threshold is placed on the dimension of subjective strength (Hahn & Harris, 2014). Then, stimuli with values that sit above the threshold receive one response (e.g., "valid" or "healthy"), whereas stimuli that sit below the threshold receive another (e.g., "invalid" or "unhealthy"). Therefore, where the decision threshold is placed is an important factor in overall cognitive performance, and can affect accuracy (% correct).

As instantiated by Stephens et al. (2018, 2020; see also Rotello & Heit, 2009), single-process signal detection models are one-dimensional (1D) – they assume only a single strength dimension, such that both seemingly "intuitive" and "deliberate" judgments (e.g., in fast vs. slow conditions) are based on a common assessment. On the other hand, dual-process models are two-dimensional (2D) – they assume that "intuitive" and "deliberate" judgments are based on two different strength dimensions, one based primarily on the output of Type 1 processing and the other based primarily on the output of Type 2 processing, respectively. Successful versions of both 1D and 2D models also include independent decision thresholds for "intuitive" and "deliberate" judgments, and hence the model variants are referred to as the *independent-1D* and *independent-2D* models (see Stephens et al., 2018).

Figure 2 applies these single-process and dual-process signal detection models to nutrition judgments, made within a two-response paradigm (with fast then slow responses). The healthiness of food products is assumed to vary along continuous dimension(s) of subjective healthiness assessment. When making nutrition judgments, an individual has to distinguish healthy products from unhealthy products based on the available cues, such as FOP labelling, branding, or the detailed nutritional information panel. According to the single-process independent-1D model in Figure. 2a, both responses in the two-response paradigm are based on a single subjective healthiness continuum. There is thus a single discriminability parameter, *d’,* for the separation between distributions of healthy and unhealthy stimuli. In Figure 2a, there are also two distinct decision thresholds or criteria parameters; one for the fast Response 1 (*C-R1*) and one for slower Response 2(*C-R2*). For example, when making fast, initial responses, the individual may place a low decision threshold, judging only the very low values as unhealthy products. When making slow and subsequent responses, however, the individual may place a relatively high threshold, so that the healthiness judgment criteria is more stringent. In contrast, the dual-process model independent-2D in Figure 2b assumes that both initial Response 1 and subsequent Response 2 judgments are based on two different strength dimensions, one based primarily on Type 1 processing and the other based more on Type 2 processing, respectively. Thus, there are two *d’* parameters, as well as the two response thresholds.

Figure 2.

Single-Process and Dual-Process Signal Detection Models of Healthiness Judgments.



Note. The continuum represents subjective healthiness assessment, and black dots indicate possible individual healthiness assessments for stimuli of differing strength. a) A single-process model (the independent-1D model), with discriminability parameter, *d'*, and decision criteria C-R1 for fast, intuitive judgments (Response 1) and C-R2 for slow, deliberate judgments (Response 2). b) A dual-process model (the indepdent-2D model) which additionally assumes that intuitive and deliberate judgments are based on two different strength dimensions.

Testing the Signal Detection Models with Signed Difference Analysis

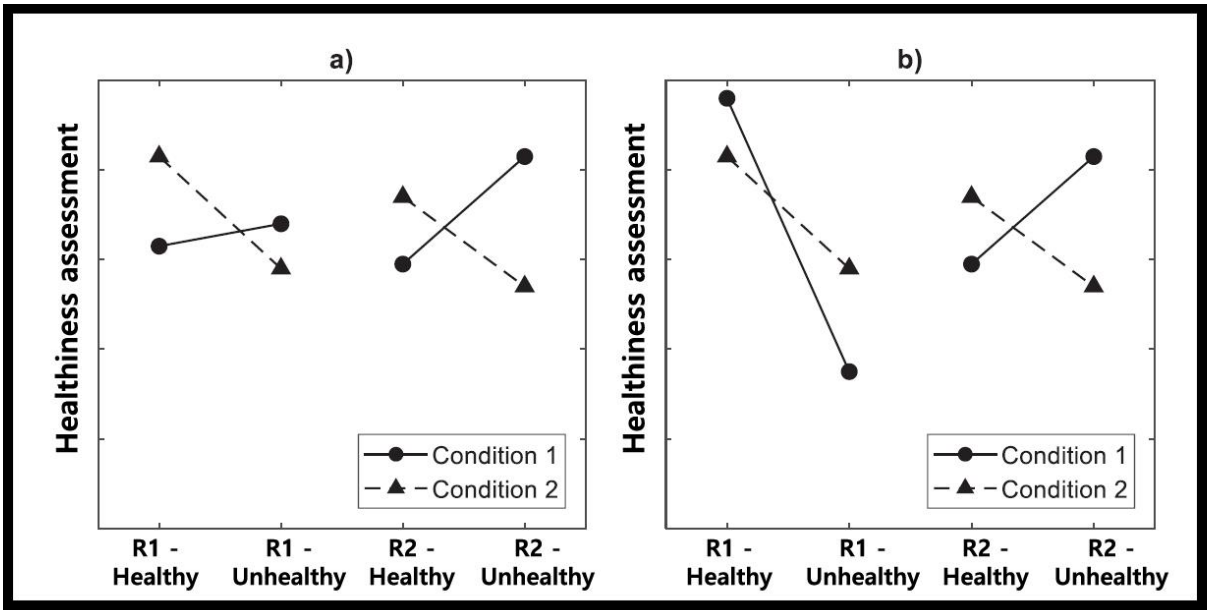
To distinguish between the competing 1D and 2D signal detection models, Stephens et al. (2018) applied a rigorous approach called Signed Difference Analysis (Dunn & James, 2003). Signed Difference Analysis allows the signal detection models to be tested in their most general form, with minimal distributional assumptions – that is, there is no commitment that the distributions of subjective strength for healthy and unhealthy stimuli are Gaussian or any other particular form. Instead, the discriminability and decision-criteria model parameters are simply assumed to have a monotonic relationship with the observed responses.

Applying Signed Difference Analysis, Stephens et al. (2018) showed that various 1D and 2D signal detection models have different "permitted" and "forbidden" ordinal data patterns. Crucially, although the independent-1D model can account for many qualitative data patterns, there is one "forbidden" ordinal pattern that it cannot account for, but the more complex independent-2D model can. Thus, a critical test of the independent-1D model is whether its forbidden pattern is observed in a two-response task with nutrition judgments.

An example of one of the many patterns permitted by both the independent-1D and -2D models is shown in Figure 3a, while an example of the qualitative pattern forbidden by the independent-1D model but permitted by the independent-2D model is shown in Figure 3b. The x-axis plots the four "dependent variables" in the Signed Difference Analysis approach, which in this case are defined as the proportion of endorsements that a food label is "healthy" for: Response 1 for a food label with a – Healthy NIP, Response 1 for a food label with an – Unhealthy NIP, Response 2 for a food label with a – Healthy NIP and Response 2 for a food label with an – Unhealthy NIP. Signed Difference Analysis involves testing observed ordinal patterns of difference between "conditions" across these four dependent variables. The one-dimensional model assumes that heuristic and intuitive judgments share a single healthiness "strength" continuum and thus one discriminability parameter. Therefore, the kind of ordinal pattern shown in Figure 3a is permitted, which is consistent with improved discrimination of healthy versus unhealthy food labels in Condition 2 compared with Condition 1, for both fast Response 1 and slow Response 2. However, as there is only one dimension of stimulus strength, across two conditions, people cannot be better at distinguishing healthy and unhealthy food products in Response 1 but simultaneously worse at distinguishing healthy and unhealthy food products in Response 2 (or vice versa). This kind of ordinal pattern is illustrated in Figure 3b (reversed cross-over patterns): relative to Condition 2, Condition 1 suggests opposing shifts in discriminability for fast Response 1 and slow Response 2. In other words, if the independent-1D model is correct, food product healthiness discrimination in the two-response task should never be found to both increase for Response 1 and decrease for Response 2 (or vice versa) across two different experimental conditions. In contrast, the independent-2D model has distinct discriminability parameters for Response 1 and 2, thus can account for such opposing shifts in healthiness discrimination.

Figure 3

Hypothetical Data Patterns in Signed Difference Analysis of the Models



Note. a) Example of a pattern that is permitted by the independent-1D model. b) Example of a pattern that is forbidden by the independent-1D model. R1 = fast Response 1; R2 = slow Response 2.

The independent-1D and -2D model have so far been tested against a large number of reasoning experiments using Signed Difference Analysis. For example, Stephens et al. (2018; 2020) examined an argument evaluation task in which participants made inductive (is the conclusion plausible?) or deductive (is the conclusion logically valid?) judgments about logical arguments. The pattern forbidden by the independent-1D model has not yet been found. Thus, it was concluded that no compelling evidence against the independent-1D model have been observed for the argument evaluation task. However, although this single-process signal detection model has been successful in accounting for judgments of verbal logical arguments, it has not yet been tested in other domains such as nutrition judgments.

Testing Single-Process and Dual-Process Models of Nutrition Judgments

As a first step towards testing the independent-1D and independent-2D models in the domain of nutrition judgments, the current study applied the two-response task that has been used by dual-process theorists in reasoning research (e.g., Bago & De Neys, 2017; Thompson & Johnson, 2014). The basic procedure for this task was explained earlier but there are other additional procedural features that could be implemented to further strengthen an experiment (e.g., impose response deadlines or add a secondary cognitive load task). However, given that the two-response task had not been applied to nutrition judgments before, and that the testing of the signal detection models was also novel in this domain, the current study was based on the simplest variant of the two-response task, as a sensible initial experiment. Modelled on Experiment 1 of Bago and De Neys (2016), the participants were instructed to give their initial response as fast as possible (Response 1), and then spend as much time as they wanted to consider food label stimuli before giving their second response (Response 2). To further enhance the distinction between Response 1 and 2, Response 1 was instructed to be based on an initial, overall impression of the food label, while Response 2 was to be focussed on the Nutrition Information Panel (these instructions are analogous to the induction versus deduction reasoning instruction used by Stephens et al., 2018, 2020).

As food label stimuli in the current experiment, three cues were included – a front-of-pack label (Health Star Rating), NIP and branding, which might differentially affect Type 1 and 2 processing under a dual-process account. Arguably, a HSR is a useful and simple heuristic cue, amendable to intuitive Type 1 processing. The HSR is a FOP rating system developed by the Australian government. The HSR assigns different stars according to the food product's healthiness, based on the energy level and the positive and negative nutrients, especially saturated fat, sugar, and salt content (Food Standards Australia New Zealand, 2018). Consumers can compare HSRs when considering similar food products, and this system has been chosen as clearly preferred label by consumers (Neal et al., 2017). One of the main reasons is that the HSR can be utilized well across groups with a range of different levels of nutritional knowledge. However, many researchers have highlighted shortcomings of HSRs – until recently (in 2020), it was based on the serving suggestion (which can be misleading), and it does not differentiate processed sugar and natural sugar (Hleborodova, 2018; Lai et al., 2019). Therefore, even though a HSR acts as a fast healthiness assessment cue, the actual healthiness of the food product can be different. Another intuitive cue, even more so than HSR, is the logo or branding of the food product. These cues' designs and wordings are entirely dependent on the marketer's intention and purpose, which usually is to appeal to consumer's needs and expectations, regardless of true healthiness value of the food product (Carrillo et al., 2012; Schneider & Pcheptsova, 2020).

The final food label cue included in the current study is the Nutrition Information Panel, which may need more Type 2 processing to be accurately evaluated. Note that in the current experiment, the NIP will be the basis of the correct, normative healthy or unhealthy response. In NIPs, each nutrient type is categorized and compared by serving size and per 100g (Eat For Health, n.d.). Listed categories include energy, protein, fat, carbohydrate, fibre and sodium, and ingredients are also listed (although ingredients were omitted in the current study). The advantage of NIPs is that consumers can consider and compare different nutrient types and make more accurate healthiness assessments. One of the biggest problems with a NIP, however, is that it is complicated, with many elements, and therefore it takes time to read and understand its content. NIPs also require certain levels of nutritional knowledge to properly understand and make use of the information. For example, each nutrient type has different recommended healthy levels (Eat For Health, n.d.) – for example, less than 3g per 100g of saturated fat is ideal, whereas food with more than 400mg per 100g of sodium level are considered unhealthy.

The aim of the current experiment was to test for differential effects of NIP, logo and HSR on healthiness judgments in a two-response task. Classic verbal dual-process accounts would predict that the more intuitive HSR and logo may have a larger effect than the NIP on the first response, while the more complex NIP should have a larger effect than the HSR and logo on the second response. However, another key aim was to then consider the results from the viewpoint of Signed Difference Analysis, and examine whether the pattern forbidden by the independent-1D model was observed, which would rule out that model in favour of the 2D model.

Methods

This experiment applied the methodologies and general design used by prominent dual-process theorists for reasoning stimuli such as logical arguments (Evans & Stanovich, 2013). In particular, the basic two-response paradigm of Bago and De Neys (2017) was implemented. However, in this experiment, participants assessed the healthiness of food label stimuli. The experiment design was fully within-participants, 2 (first vs. second rating response) x 2 (healthy vs. unhealthy NIP) x 2 (healthy vs. unhealthy HSR) x 2 (healthy vs. unhealthy brand/logo).

Participants

A total of 60 participants were tested (15 males, 45 females, Mean age = 19.67, *SD* = 2.16). Most of the participants were recruited from the first-year Psychology participant pool at the University of Adelaide and received course credit for participation, at the standard rate, except for two participants who were recruited from the general community. Only participants who are fluent in English, have normal or corrected vision, and have no current neurological, learning or intellectual disorder were allowed to participate in the study. Six possible outlier participants were identified (e.g., based on uncommon responses to some stimuli), but the conclusions were unchanged if they were excluded. Therefore, all participants were retained for data analysis.

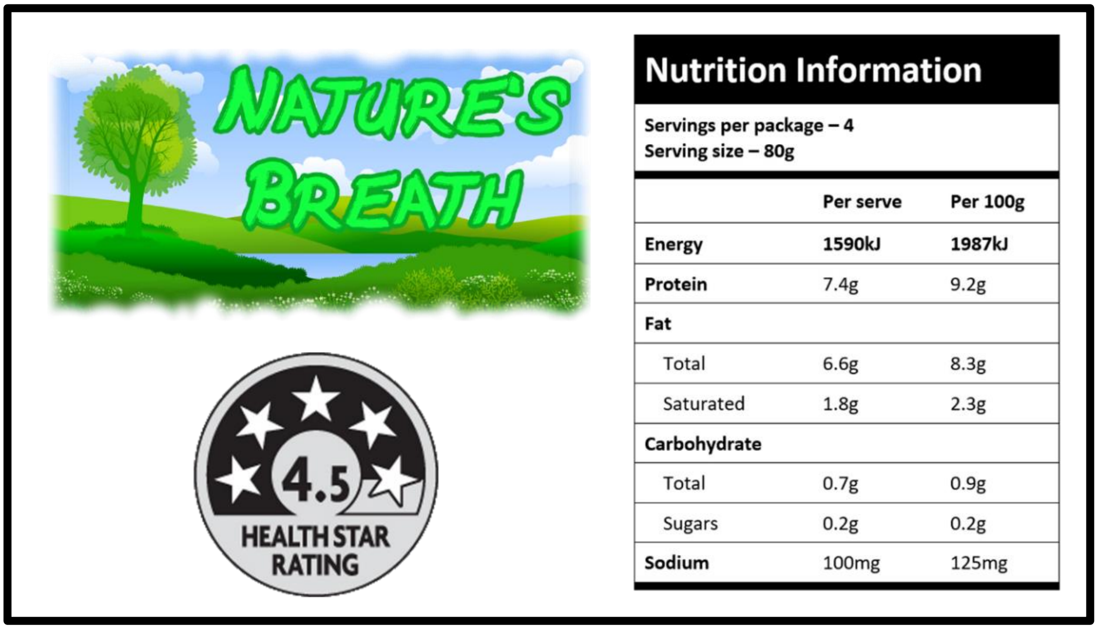
Materials

Participants were presented with 36 trials, each showing the labels of fictional food products (see Figure 4 for an example). The provided food products showed FOP nutritional label (Health Star Rating), fictional product logo and detailed Nutrition Information Panel. The main 32 trials were formed by factorially crossing three factors: HSR (healthy vs. unhealthy), NIPs (four healthy and four unhealthy), and logos (healthy vs. unhealthy). Also, four filler trials composed of a neutral HSR, neutral nutritional panel and neutral fictional brand name were created.

In this experiment, a HSR of 1 was classified as being ‘unhealthy’, 3.5 as being ‘neutral’ (for the filler trials) and 4.5 as being ‘healthy’. The HSR images used for designing the food labels were downloaded from the government website for the Health Star Rating System (Food Standards Australia New Zealand, 2018).

Figure 4

Example of the labels of fictional food products composed of logo, HSR and NIP



	Per serve	Per 100g
Energy	**1590kJ**	**1987kJ**
Protein	7.4g	9.2g
Fat		
Total	6.6g	8.3g
Saturated	1.8g	2.3g
Carbohydrate		
Total	0.7g	0.9g
Sugars	0.2g	0.2g
Sodium	100mg	125mg

The NIP design was based on the Australian government guideline (Eat For Health, n.d.). Amongst the nutrient values contained in the standard NIP, it was decided that energy, saturated fat, sugars and sodium would be varied in this experiment, to form the four unhealthy and four healthy NIPs. This was because these four key nutrient values affect real HSR. As shown in Table 1 below, across the four key nutrients, the healthy NIPs each had three low values and one high value, while the unhealthy NIPs had three high values and one low value. For example, Healthy NIP 1 had a high level of energy but a low level of saturated fat, sugars and sodium, whereas Healthy NIP 2 had a high level of saturated fat but the other three nutrients were of a low level.

Table 1

The Structure Used for the Nutrition Information Panels

Panel	Energy	Saturated Fat	Sugar	Sodium
Healthy NIP 1	High	Low	Low	Low
Healthy NIP 2	Low	High	Low	Low
Healthy NIP 3	Low	Low	High	Low
Healthy NIP 4	Low	Low	Low	High
Unhealthy NIP 1	Low	High	High	High
Unhealthy NIP 2	High	Low	High	High
Unhealthy NIP 3	High	High	Low	High
Unhealthy NIP 4	High	High	High	Low
Neutral	Neutral	Neutral	Neutral	Neutral

Each NIP had different values for the four key nutrients, and similar values for the remaining nutrients (see Table 2 and Figure 4). The key nutrient values were decided based on two sources: *How To Understand Food Labels*, distributed by Department of Health and Ageing (Eat For Health, n.d.), and the *Guide for industry to the Health Star Rating Calculator (HSRC)* (Food Standards Australia New Zealand, 2018). The first source outlines the recommended value for each nutrient, to guide healthy eating, which was used as the standard (or boundary) that classifies healthy and unhealthy values. Then, the actual value for each nutrient was decided by referencing HSR "baseline points" from the HSR Calculator. The higher the HSR baseline points are, the unhealthier the food product is, and therefore healthy nutrient value was set lower, and unhealthy nutrient value higher, than the standard value. For example, *How To Understand Food Labels* recommends less than 3g of saturated fat per 100g, so 3g per 100g became the standard nutrient value for saturated fat. This value of 3g per 100g of saturated fat is scored as baseline points of 3 in the HSR Calculator. Therefore, nutrition values around 1.0g per 100g, which scores 1 baseline point, were defined as healthy, and 5.0g per 100g, which scores 5, as unhealthy for saturated fats. As for the neutral Nutrient Information Panels for filler trials, the nutrient values were in line with the standard values. For all NIPs, protein was around 9g, total fat was around four times greater than saturated fat, and total carbohydrate was around five times greater than sugars. The serving sizes were all around 85g, and the "per serve" values were calculated accordingly.

Table 2

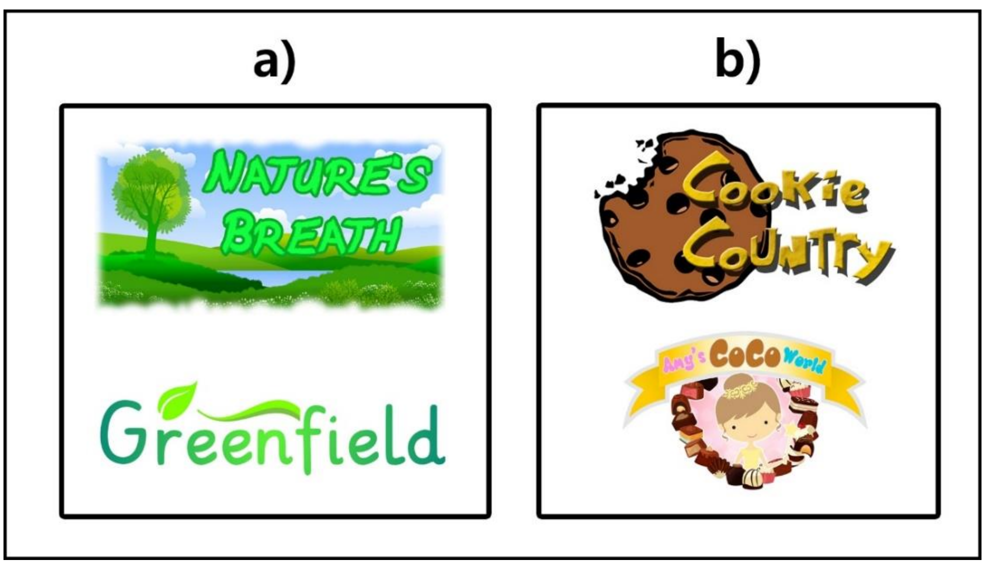
Critical Nutrient Values Used in Each NIP (Per 100g)

Panel	Energy	Saturated Fat	Sugar	Sodium
Healthy NIP 1	1987kJ	2.3g	0.2g	125mg
Healthy NIP 2	987kJ	8.1g	6.2g	50mg
Healthy NIP 3	950kJ	2.1g	27.9g	32mg
Healthy NIP 4	835kJ	1.1g	3.7g	810mg
Unhealthy NIP 1	910kJ	5.2g	26.5g	454mg
Unhealthy NIP 2	1700kJ	1.8g	32g	710mg
Unhealthy NIP 3	2058kJ	8.6g	8.7g	964mg
Unhealthy NIP 4	2170kJ	15.1g	44.9g	161mg
Neutral NIP 1	1350kJ	2.8g	17.9g	440mg
Neutral NIP 2	1385kJ	3.2g	19.2g	465mg

Fictional food brand logos also generated. Initially, 18 logos were created, 8 of which were suggestive of a healthy brand and the other 8 of an unhealthy brand. Perceived healthiness was assessed via an initial online survey in which participants were asked to assess how unhealthy or healthy each logo appears. A total of 34 participants were tested (7 males, 27 females, Mean age = 30.42, SD = 10.08). The logos were presented one at a time in a random order and each participant chose one of the following ratings: highly unhealthy, moderately unhealthy, slightly unhealthy, slightly healthy, moderately healthy and highly healthy. The data were then coded to 1, 2, 3, 4, 5 and 6 respectively and mean and standard deviation for each logo was calculated. The two logos that participants rated the unhealthiest (M = 1.85, SD = 0.70; M = 1.76, SD = 0.89) and the two healthiest (M = 5.09, SD = 1.11; M = 5.09, SD = 0.97) were used in the main experiment (see Figure 5). In the main experiment, the two healthy logos were treated as equivalent, as were the two unhealthy logos, and they were randomly assigned to the food labels including the fillers with neutral NIPs.

Figure 5

Fictional Food Brand Logos Used in The Experiment

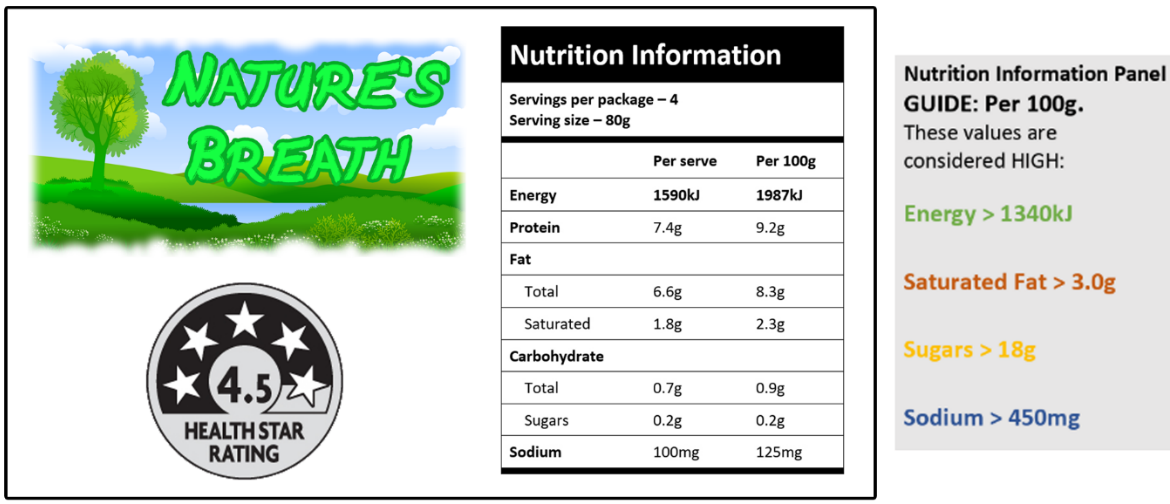


Note. a) Two logos that participants rated the healthiest. b) Two logos that participants rated the unhealthiest.

Throughout the experiment, each food label was paired with a guide to help participants with interpreting the complex Nutritional Information Panel, as shown in Figure 6. The guide was based on *How To Understand Food Labels* (Eat For Health, n.d.).

Figure 6

Example of a Food Label Paired with a Guide to Interpreting the Nutrition Information Panel.



Procedure

The experiment was run online using Qualtrics. Participants were informed that this research project was being conducted to investigate how people assess the healthiness of food products, based on the information provided on food packaging. They were clearly instructed that the label of a fictional food product would be shown on each page, and their job was to rate the healthiness of the food. They were shown an example food label, depicting all three components: a logo of the company that makes the food, a Health Star Rating and a Nutrition

Information Panel. Simple explanations of the Health Star Rating and Nutrition Information Panel and how to interpret them were also provided (see Appendix A). Participants were asked to assess the healthiness of different food products twice; the first rating was to be based on their initial impression, and the second rating was to be made after careful examination of the NIP. The literal instructions that were used, stated the following:

"For each food label, please make two ratings:

RATING 1:

Quickly rate how unhealthy or healthy the food appears, based on your initial, intuitive impression of the entire label. We want you to respond with the very first rating that comes to mind, as quickly as possible.

RATING 2:

Then, the label will be shown again. Take all the time you want to carefully examine the Nutrition Information Panel, and rate the food's healthiness based on this Panel.

Use the Guide shown on the right of the screen, to help you interpret the Nutrition Information Panel."

During the experiment task, all participants were presented with one food label at a time, shown in a randomly determined order. Each trial was comprised of two pages. The first page depicted an image of the label of a fictional food product with a guide for interpreting the Nutrition Information Panel. Underneath the image was the instruction to rate the food's healthiness based on initial impression of the entire label. The literal instruction was as follows:

"RATING 1: As quickly as possible, rate how unhealthy or healthy this food product is, based on your initial impression of the entire label."

The participants had to choose one of the following ratings: Highly Unhealthy, Moderately Unhealthy, Slightly Unhealthy, Slightly Healthy, Moderately Healthy and Highly Healthy. After they rated their initial impression and clicked the Next button, the second page was shown. The second page depicted the same image and underneath was the instruction to rate the food's healthiness based on examination of the Nutritional Information Panel. The literal instruction was as follows:

> "RATING 2: Take all the time you want – rate how unhealthy or healthy this food product is, based on careful examination of the Nutritional Information Panel."

The same rating scale was presented again. After participants completed the second rating and clicked the Next button, the next trial was shown. Participants could not return to previous pages.

Results

The following section presents analysis of variance (ANOVA) tests on ratings from Response 1, where the participants were instructed to give their fast, initial, overall impression, and for Response 2, where the participants were instructed to take their time and focus on the NIP. The effects of HSR, NIP and Logo on the two responses will be examined and then compared. A subsequent section examines whether the pattern forbidden by the independent-1D model has been observed.

Before proceeding, the mean response time measured for each condition was examined and compared. Appendix B shows that, even though the differences in some cases were small, the mean response time for each condition was indeed faster in Response 1 than Response 2. This suggests that participants generally followed the different instructions for Response 1 and 2, especially considering the fact that ratings made for Response 2 could have been relatively quick, since they were based on viewing the food label across both pages, for Response 1 and 2.

The Effects of Food Label Features on Response 1 and Response 2

Three-way repeated measures ANOVA was performed to evaluate the effects of healthy versus unhealthy HSR, Logo and NIP on the healthiness assessments of food products. Initially, Response 1 and 2 were examined in separate tests.

The ANOVA output for Response 1 (Table 3; Figure 7 and 8) shows that there were statistically significant effects of HSR ($p = <.001$, ges = 0.567), Logo ($p = < 0.001$, ges = 0.133) and NIP ($p = 0.003$, ges = 0.014) on the healthiness assessment, with higher ratings for healthy than unhealthy features. According to the *ges* values, a large proportion of variance was accounted for by HSR and Logo, but only a small proportion of variance was accounted for by NIP. As the HSR effect had the highest *ges* value of 0.567, HSR was the most influential factor when making fast and intuitive judgment. Table 1 also shows that there were neither statistically significant two-way interactions nor significant three-way interactions between HSR, Logo and NIP.

Table 3

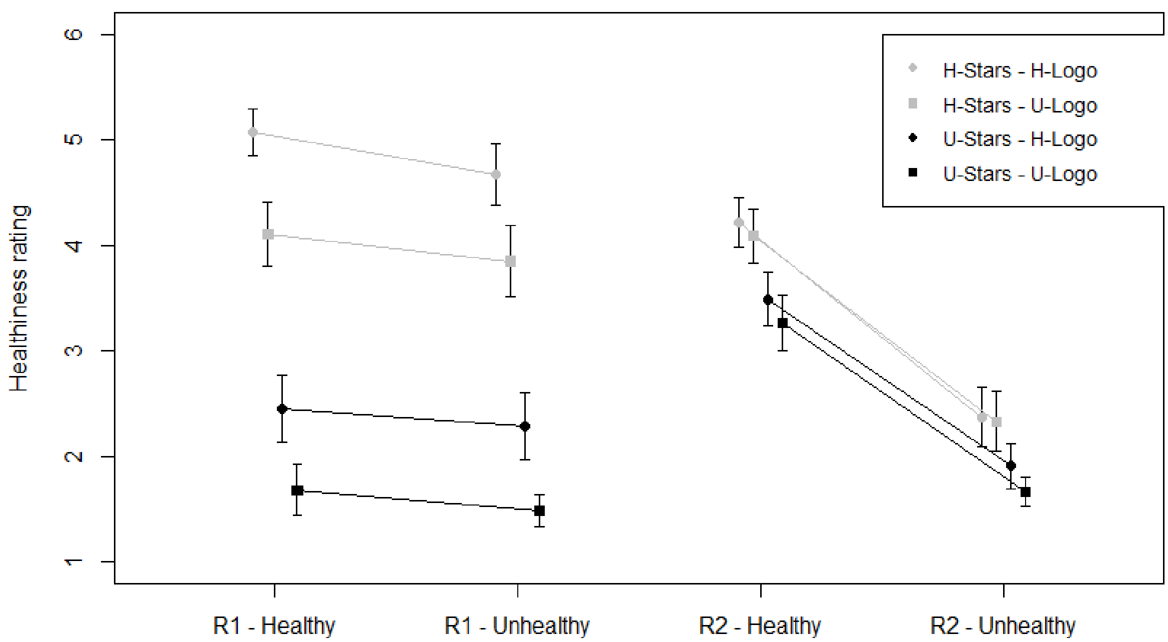
Three-way Repeated Measures ANOVA Results for Response 1

Effect	DFn	DFd	*F*	*P*	*ges*
HSR	1	59	170.075	< 0.001*	0.567
Logo	1	59	63.263	< 0.001*	0.133
NIP	1	59	9.943	0.003*	0.014
HSR x Logo	1	59	1.638	0.206	0.001
HSR x NIP	1	59	3.597	0.063	0.001
Logo x NIP	1	59	0.738	0.394	0.000
HSR x Logo x NIP	1	59	1.072	0.305	0.000

Note. * = $p < 0.05$; *ges* = "generalised eta squared".

Figure 7

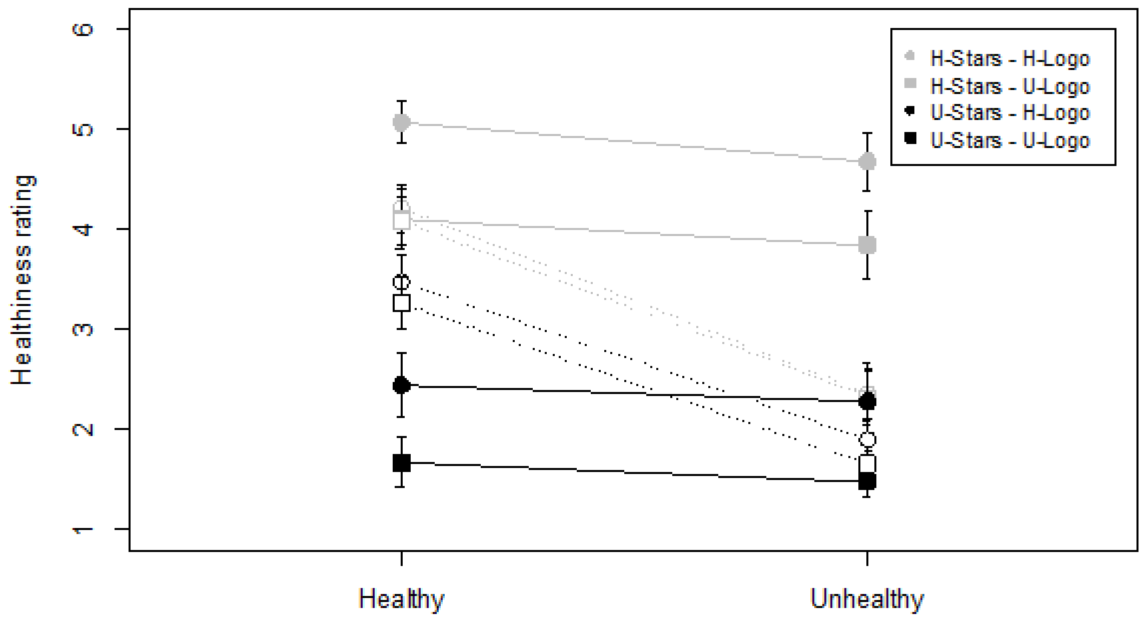
Healthiness Assessment of Food Products by Participants



Note. R1 = Response 1, R2 = Response 2, Healthy = Healthy NIP, Unhealthy = Unhealthy NIP, Stars = HSR, H = Healthy, U = Unhealthy. Error bars show 95% confidence intervals.

Figure 8

Healthiness Assessment of Food Products by Participants (Response 1 and Response 2 Plotted Together)



Note. Response 1 is signalled by the solid lines and filled marks; Response 2 is signalled by the dashed lines and unfilled markers.

A second ANOVA was performed for Response 2. Similar to the results for Response 1, Table 4 and Figure 7 show that there were statistically significant effects of HSR ($p < 0.001$, ges = 0.112), Logo ($p < 0.001$, ges = 0.007) and NIP ($p < 0.001$, ges = 0.448) on the healthiness assessment, with higher ratings for healthy than unhealthy features. Unlike Response 1, however, there was a larger effect of NIP and a smaller effect of Logo and HSR. Therefore, it can be said that Response 2 was most strongly influenced by the NIP. Table 4 also shows that, again, there were neither statistically significant two-way interactions nor significant three-way interactions between HSR, Logo and NIP for Response 2.

Table 4

Three-way Repeated Measures ANOVA Results for Response 2

Effect	DFn	DFd	*F*	*p*	*ges*
HSR	1	59	41.417	< 0.001*	0.112
Logo	1	59	13.598	< 0.001*	0.007
NIP	1	59	232.161	< 0.001*	0.448
HSR x Logo	1	59	3.124	0.082	0.002
HSR x NIP	1	59	3.703	0.059	0.0003
Logo x NIP	1	59	0.172	0.679	0.000
HSR x Logo x NIP	1	59	0.871	0.354	0.000

Note. * = $p < 0.05$; ges = “generalised eta squared”.

To test whether there were significant *differential* effects of HSR, NIP and Logo on Response 1 versus Response 2, a four-way repeated measures ANOVA was also performed. In addition to HSR, Logo and NIP, the effect of Response Type (Response 1 or Response 2) on the participants’ healthiness assessment was evaluated. Appendix C presents all the results from this test, but of core interest was whether there were significant two-way interactions between Response Type and each of HSR, NIP and Logo. Indeed, there were statistically significant interactions between Response Type and each of the other three factors: HSR ($F(1,59) = 83.45$, $p < .001$, $ges = 0.163$), Logo ($F(1, 59) = 37.65$, $p < .001$, $ges = 0.028$) and NIP ($F(1, 59) = 155.0$, $p < .001$, $ges = 0.113$). This supports that there were larger effects of HSR and Logo for Response

1 than for Response 2, but a larger effect of NIP for Response 2 than for Response 1 (see Figure 7).

It is important to note that care must be taken in interpreting interaction effects identified by ANOVA as supporting that there are differential effects on the underlying latent psychological process(es), such as people's subjective healthiness assessment (see Loftus, 1978; Wagenmakers et al., 2012). The key issue is that ANOVA does not distinguish "removable interactions" (interactions that can be undone by a monotonic transformation of the observed responses) from non-removable interactions, which offer more rigorous support for differential effects on latent processes. However, in this case there *were* some non-removable, cross-over interactions, particularly for some of the conditions with unhealthy HSRs. This can be seen more clearly in Figure 8, where healthiness assessment for Response 1 was plotted directly over healthiness assessment for Response 2. Therefore, there are clear, differential effects of the food label features on people's initial intuitive healthiness assessments versus the second, more considered assessments. This finding is consistent with the dual-process view that there are two distinct psychological dimensions of assessment strength. However, it may also be consistent with a single-process model with multiple parameters, such as the independent-1D model.

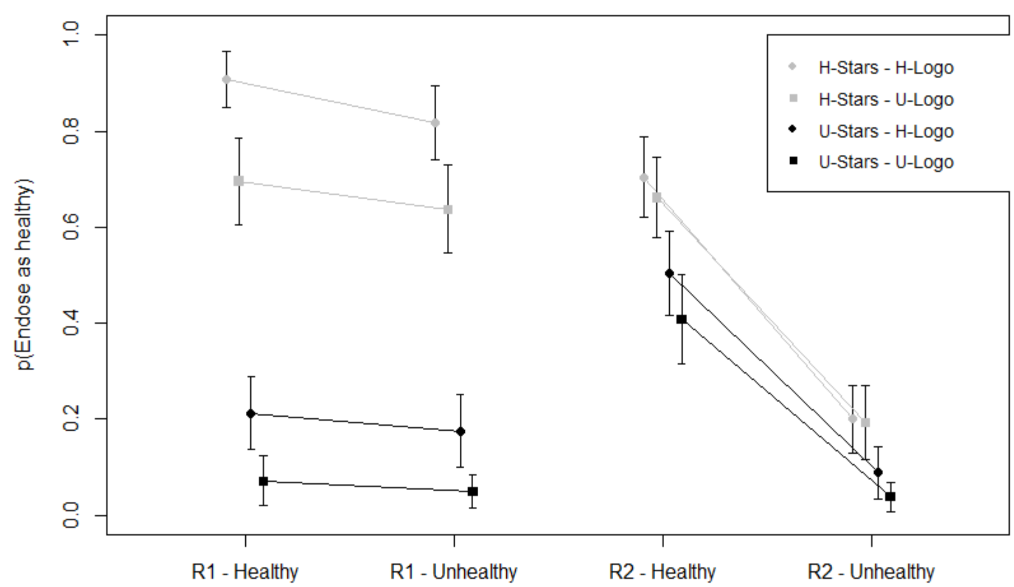
Signed Difference Analysis Examination of the Independent-1D Model

Lastly, the results were considered from the viewpoint of Signed Difference Analysis to test the single- and dual-process signal detection models in their most general form, with minimal distributional assumptions. The goal was to examine whether the ordinal pattern that is forbidden by the independent-1D model was observed; this would rule out that model in favour of the independent-2D model. The models and current Signed Difference Analysis approach (see Stephens et al., 2018) are based on binary judgments, so first the 6-point healthiness ratings were converted to binary responses. Ratings of 3 ("slightly healthy") or higher were counted as an endorsement that the food label was healthy.

Figure 9 shows the mean proportion of endorsements, plotted according to the design for the Signed Difference Analysis, with the four "dependent variables" on the x-axis, and the conditions as different lines. Similar effects of NIP, Logo and HSR were observed for binary responses as in Figure 7 with the mean ratings. Crucially, the figure does not reveal any ordinal patterns corresponding to the double cross-over pattern that is forbidden by the independent-1D model (cf. Figure 3b). Therefore, despite the differential effects of HSR, Logo and NIP on Response 1 versus 2 (as identified by the four-way ANOVA), there was no compelling evidence against the single-process, independent-1D model.

Figure 9

Proportion of Healthiness Endorsement of Food Products by Participants



Note. R1 = Response 1, R2 = Response 2, Healthy = Healthy NIP, Unhealthy = Unhealthy NIP, Stars = HSR, H = Healthy, U = Unhealthy. Error bars show 95% confidence intervals.

Discussion

Summary of Overall Findings

There were two main aims for this two-response paradigm experiment. The first aim was to test for differential effects of NIP, logo and HSR on people's initial, "intuitive" healthiness judgments of food labels versus slower, subsequent judgments that were more focused on the NIPs. According to the classic dual-process theories of cognition (Bago & De Neys, 2017; Sanjari et al., 2017), fast and heuristic Type 1 processing will be more affected by the HSR and logo, whereas slow and analytical Type 2 processing will be more affected by the NIP. In turn, more Type 1 processing should be involved with the initial Response 1, whereas more Type 2 processing should be involved with Response 2. The second aim of this experiment was to investigate whether there was evidence against a single-process signal detection model, in favour of a dual-process alternative.

To address these aims, a two-response task was implemented. Each question showed a fictional food product label composed of three cues (HSR, logo and HSR) that varied according to their healthiness. Participants first rated the food product's healthiness based on their initial impression of the entire label, as quickly as possible. Participants then made a second healthiness rating based on careful examination of the NIP.

The results show that the three food label cues all had an effect on the participants' healthiness judgments of food products. Moreover, differential effects of HSR, Logo and NIP on Response 1 versus 2 were observed, with the first two cues having a bigger effect on Response 1, but the NIP having a bigger effect on Response 2, as predicted by classic dual-process theory. However, when the results were considered from the viewpoint of Signed Difference Analysis of the formal independent-1D and independent-2D models, the ordinal pattern forbidden by the independent-1D model was not observed. Thus, the single-process independent-1D model can be retained.

Discussion of Findings

The results supported the expected findings that the HSR and logo would have a bigger effect on initial fast judgments and the NIP would have a bigger effect on the consequent slower judgments. On the one hand, these differential effects are surprising given that previous reasoning studies involving the two-response task have found that people often report the same judgment for Response 1 and 2 (e.g., Thompson et al, 2011). On the other hand, these findings are sensible given the design and nature of the food stimuli. For example, the HSR presents nutritional information in an easily understandable format; the concept of higher star ratings reflecting better quality of the product is familiar and thus it is a more salient cue compared to reading and comprehending the detailed numerical values of the NIP. Similarly, the logo presents salient visual cues which suggest the nutritional status of the products. Even though both the HSR and logo may not correctly reflect the true nutritional information, these are more readily available cues for the participants to read and understand in a short period of time.

Amongst the differential effects of the food label cues, the most interesting result is that of the HSR. In Response 1, the HSR had a bigger effect compared to the logo, which suggests that the participants found the HSR to be either more reliable or a more easily comprehendible source of information. This could be because the star rating system provides clearer basis for judgment, and perhaps also because the participants were unfamiliar with the provided fictional logos. Another interesting result of the HSR is that, unlike the logo, it still had a substantial effect on Response 2 as well. Thus, participants continued to use this cue even when asked to focus on the NIP for Response 2. However, the experiment was necessarily designed so that the participants were asked to give their answer for Response 1 and then immediately to give Response 2 for the same food label. Thus, it could be possible that the HSR viewed for Response 1 had a lingering effect on the participants' judgment during Response 2.

Another purpose of this experiment was to more rigorously examine evidence of dual-process theories. This two-response task was designed in accords with methodologies used by prominent dual-process theorists, under the assumption that Response 1 reflects mostly fast and heuristic Type 1 processing and Response 2 reflects more slow and intuitive Type 2 processing. The differential effects of food label features on each response are – consistent with this view.

However, dual-process theories are not the only possible explanation. Previous studies have criticized the validity of dual-process theories and suggested single-process theories as an alternative framework to account for assessments and judgments (Kruglanski & Gigerenzer, 2011; Stephens et al., 2018, 2020). In the current experiment, the signal detection models and Signed Difference Analysis approach of Stephens et al. (2018) were implemented to formally test the competing theories. The key result was that no compelling evidence against the single-process independent-1D model was found. This model assumes that a distinct decision threshold is placed for Response 1 and Response 2, along a single underlying continuous dimension of healthiness.

Note that this does not refute dual-process theories in favour of single-process theories. What the success of the 1D model represents is support for this single-process model as a viable alternative to the dual-process independent-2D model. In other words, both models of healthiness judgment can account for the results of this experiment. Nevertheless, the results indicate that, despite the current status of dual-process theories as a widespread and influential cognitive framework, the necessity of dual-process explanations should be reconsidered. Instead, comparatively simpler and equally successful single-process theories need to be considered as well.

Theoretical and Practical Implications

The most theoretically significant aspect of this experiment is the rigorous comparison of competing single- and dual-process theories of nutrition judgments. This project demonstrates how the theories can be formally instantiated using signal detection models such as the independent-1D and -2D models, drawing on models that have previously been applied to logical reasoning tasks (Stephens et al., 2018). Results from these reasoning tasks have consistently shown that a wide range of judgments can be explained with the quantitative, single-process, independent-1D model. However, since this reasoning evidence was from a different domain of cognition, the success of the independent-1D model may not have generalised to nutrition judgments. Therefore, examination of the results from nutrition judgments made in the two-response task is important as this was the first study to test the single-process signal detection model in another domain. As no compelling evidence against the independent-1D model was observed from the results, this study further supports the view that single-process theories can also account for cognitive assessments and judgments and can be considered as a viable alternative to dual-process theories.

Therefore, the results from this study have a broad range of implications across many different areas of psychology. Dual-process theories have become an appealing framework within cognitive psychology, rapidly increasing in popularity. Examples include explicit versus non-explicit memory and learning (Schacter & Tulving, 1994; Squire, 1992), declarative rule-based versus procedural category learning (Ashby et al., 1998), automatic versus explicit processing of social information (Evans, 2008; Smith & DeCoster, 2000), multiple processes for utilitarian versus deontological moral judgments (Paxton & Greene, 2010), holistic and featural processing of faces (Tanaka & Gordon, 2011) and visuo-spatial versus phonological working memory (Baddeley, 2012). Even previous studies regarding nutrition label formats and healthiness judgment (Sanjari et al., 2017), the topic of this experiment, have applied dual-process theories. However, important evidence for such accounts is often faced with criticisms such as its basis on functional dissociations (Stephens et al., 2020), unreliable distinction between two types of processing (Osman, 2004), and its structure as a generic framework rather than a unified theory (Grayot, 2019). In response to the widespread use of dual-process theories despite such major deficiencies, there has been a growing number of claims that single-process theories have been prematurely disregarded in several different areas. These include category learning (Newell et al., 2010; Stephens & Kalish, 2018), face perception (Loftus et al., 2004) and recognition memory (Dunn, 2008; Hayes et al., 2017), and now nutrition judgment is added to this list. As in these other domains, the current study similarly performed targeted experiments to see whether single-process models really can be rejected, and the result aligned with the findings from other experiments that formal single-process models cannot be ruled out (cf. Stephens et al., 2020). This research could act as another example for future studies of single-process theories and more elaborate and meticulous comparison of competing single- and dual-process theories could be further developed and applied in many different domains of psychology. If such approaches are applied across areas where dual-process theories have been – perhaps unquestionably – used as a model of cognition, and produce similar results to the current study, then the argument that single-process theories are a reasonable alternative explanation to dual-process theories would be more widely considered and accepted.

The results from this study also have potential for important practical implications in everyday life. The main focal point from this experiment's results is the influence of different food label cues on people's healthiness assessment of the food products. Several studies have examined the effectiveness of different types of food labels, to identify the best method to provide customers with accurate nutritional information, in a more easily comprehendible format (Neal et al., 2017). According to World Health Organization report (2019), seventy percent of deaths worldwide were due to non-communicable diseases (NCDs), such as type 2 diabetes, cancer, and cardiovascular disease. Crucially, unhealthy diets that provide insufficient or excessive amounts of energy, nutrients and other components, which could possibly lead to obesity, was listed as one of the five biggest risk factors. were reported to be obese in 2017-18, which marks an increase from 19% in 1995 (Australian Institute of Health and Welfare, 2019). Therefore, it is of great importance to provide consumers with accurate dietary information about food products so it may assist them to make appropriate food choice for their health.

The differential effects of the food label cues used in the current study provide relevant information for practical use. Since the results show a large effect of the HSR on Response 1 and a smaller but still reliable effect on Response 2, this FOP rating system can be used as a key feature in assisting customers' healthiness assessment of food products in real-life environment. Since the HSR was established in 2014 as a preventative measure for overweight problems of Australians (Health Star Rating System, 2019), validity of the HSR has been closely monitored and evaluated by government agencies and academic researchers alike. The results from this study add to the widely recognised claims that the HSR is a fast and effective healthiness assessment cue. On the other hand, this means that when the HSR displays inaccurate or inappropriate nutritional information, people are likely to rely on it and make the wrong healthiness judgment. As evidenced by the HSR effect on the Response 2 results, after consumers have read or comprehended the HSR, they are likely to be influenced by this information even when considering detailed numerical information in the NIP. Indeed, many shortcomings of the HSR have already been highlighted (cf. Hleborodova, 2018; Lai et al., 2019), and many recommendations to improve the rating system were proposed and implemented.

Finally, the findings of the current experiment that the logo and NIP also have effects on healthiness judgments of food products aligns with previous studies (Carrillo et al., 2012; Schneider & Pcheptsova, 2020). Considering the NIP, the results from the novel two-response task in the current study indicate that, even though the NIP provides the most accurate information, it is most effective when consumers have sufficient time and encouragement to read and comprehend its complex numerical values. Therefore, simple summary guideline such as the guide used during this experiment to help customers with interpreting the NIP could be recommended, for instance on supermarket shelves or online shopping webpages. However, it should be noted that although the evaluation and potential improvement of the FOP labels could have an effect on people's nutritional judgments, this may not necessarily translate to their actual food consumption decisions or purchasing behaviour. In order to prevent or rectify health problems caused by unhealthy dietary choices, people must have the motivation and means to change their behaviour first. Only then, the findings of studies such as this will be put to practical use.

Limitations and Future Research Directions

This research may be subject to a number of limitations. One potential limitation to the generalization of the findings is that the experiment was conducted via online only. Due to state-wide social distancing measures during the period of COVID-19 alert, the experiment was unable to be conducted in the ideal laboratory conditions. Hence, the experiment was conducted online, making it difficult to ensure the participants were not distracted and that they fully understood and followed the instructions. For example, when assessing the healthiness of the food label for Response 1, instead of giving a fast response based on an initial impression, the participants may have made more considered assessments. As for Response 2, they may not have considered only the NIP as instructed, but also deliberately considered the HSR and logo as well. It is possible that the HSR's lingering effect on Response 2 could be due to this limitation. However, given the large differences in ratings given in Response 1 versus Response 2, and the expected directional effects of NIP, logo and HSR, this potential limitation does not seem to be a critical issue.

Another potential issue is the representativeness of the sample. Most participants were first year Psychology students, with an uneven gender ratio (15 males, 45 females) and narrow range of age (mean age = 19.67, *SD* = 2.16) and occupation need to be improved. This student sample may not be representative of the general population in terms of potentially relevant factors such as nutritional knowledge and interest, and health motivations. However, the current study at least attempted to ensure that all participants had an understanding of the NIPs and HSRs, by describing these cues at the outset, and including the NIP guide during the experiment. The NIP guide was included to maximise the opportunity for observing an effect of NIPs, but future research can examine healthiness ratings without such a guide.

Despite these limitations, this experiment still represents an importance advance, as it is the first step towards testing the 1D and 2D models in the domain of nutritional judgments. Therefore, conducting further experiments based on this research would be an appropriate next step. In addition to addressing the limitations listed above, a range of additional procedures used by prominent researchers need to be implemented. For example, by enforcing a strict and challenging response deadline during Response 1, experimenters can avoid overly extended response times for the fast, initial response. This may increase the opportunity for observing differential effects of Type 1 and 2 processing. Adjustments to materials could also be suggested. For instance, food labels from real products rather than fictional ones could be used as the stimuli, to simulate real life as closely as possible.

Beyond testing the competing single- and dual-process theories, training procedures could also be implemented to show that people can be trained in a short amount of time to accurately and quickly make healthiness assessments. Future studies in this direction could distinguish the most efficient means of training, which would be a useful resource in public health applications for improving people's nutrition knowledge and ability to comprehend NIPs.

Conclusion

The aim of this study was to investigate the effects of different label features on people's healthiness assessments of food products, and to examine and compare competing single- and dual-process theories. The two-response paradigm was implemented to attempt to capture Type 1 and Type 2 processing under a dual-process account. Results showed that HSRs and logos have a larger effect on fast, initial responses, whereas NIPs have a larger effect on slower, analytical second responses. This is consistent with classic dual-process theories of cognition, which state that people are more likely to make an appeal-based choice under Type 1 processing and a reason-based choice under Type 2 processing. However, when the data were examined using a Signed Difference Analysis approach, they did not show any ordinal patterns forbidden by the independent-1D model, which means no compelling evidence against single-process theories was found. This further supports the increasing body of claims that single-process accounts are a viable alternative to popular dual-process accounts. The results from this study have a broad range of implications across many different areas. Theoretically, this study supports that single-process theories have been prematurely disregarded, and practically, the differential effects of HSRs, logos and NIPs on fast versus slow healthiness assessments have been empirically demonstrated. This experiment acts as the first step towards more rigorous comparisons of competing single- and dual-process theories in the domain of nutrition judgments, and thus towards a greater understanding of the cognitive mechanisms that drive these judgments.

References

Aschemann-Witzel, J., Grunert, K.G., van Trjip, H.C.M., Bialkova, S., Raats, M.M., Hodgkins, C., Wasowicz-Kirylo, G., & Koenigstorfer, J. (2013). Effects of nutrition label format and product assortment on the healthfulness of food choice. *Appetite*, *71*. http://doi.org/10.1016/j.appet.2013.07.004

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442-481. https://doi.org/10.1037/0033-295X.105.3.442

Australian Government Department of Health. (2013, October 08). *Food Labelling*. Retrieved from https://www1.health.gov.au/internet/publications/publishing.nsf/Content/canteen-mgr-tr1~food-labelling

Australian Institute of Health and Welfare. (2019). *Overweight and obesity: an interactive insight.* Retrieved from https://www.aihw.gov.au/reports/overweight-obesity/overweight-and-obesity-an-interactive-insight

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29. https://doi.org/10.1146/annurev-psych-120710-10042

Bago, B. & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90-109. https://doi.org/10.1016/j.cognition.2016.10.014

Balasubramanian, S.K. & Cole, C. (2002). Consumers' search and use of nutrition information: The challenge and promise of the nutrition labelling and education act. *Journal of Marketing*, *66*(3), 112-127. https://doi.org/10.1509/jmkg.66.3.112.18502

Carrillo, E., Varela, P., & Fiszman, S. (2012). Effects of food package information and sensory characteristics on the perception of healthiness and the acceptability of enriched biscuits. *Food Research International*, *48*, 209-216. https://doi.org/10.1016/j.foodres.2012.03.016

De Neys, W. (2006). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *Quarterly journal of experimental psychology*, *59*(6), 1070-1100. https://doi.org/10.1080/02724980543000123

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science, 7*(1). https://doi.org/10.1177/1745691611429354

Dhar, R., & Gorlin, M. (2014). A dual-system framework to understand preference construction processes in choice. *Journal of Consumer Psychology*, *23*(4), 528-542. https://doi.org/10.1016/j.jcps.2013.02.002.

Drescher, L. S., Roosen, J., & Marette, S. (2014). The effects of traffic light labels and involvement on consumer choices for food and financial products. I*nternational Journal of Consumer Studies, 38*(3). https://doi.org/10.1111/ijcs.12086

Dunn, J. C. (2008). The dimensionality of the remember-know task: A state-trace analysis. *Psychological Review*, *115*, 426-446. https://doi.org/10.1037/0033-295X.115.2.426

Eat For Health. (n.d.). *How to understand food labels*. Retrieved from: https://www.eatforhealth.gov.au/eating-well/how-understand-food-labels

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*(8), 709-724. https://doi.org/10.1037/0003-066X.49.8.700

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255-278. https://doi.org/10.1146/annurev.psych.59.103006.093629

FitzSimons, P. (2017, June 13). '*It's freaking hopeless': why the Health Star Rating System has to go.* Sydney Morning Herald. Retrieved from https://www.smh.com.au/opinion/its-freaking-hopeless-the-health-star-rating-system-has-to-go-20170611-gwp0mb.html

Food Standards Australia New Zealand. (2018). *Guide for Industry to the Health Star Rating Calculator (HSRC) (6th ed.)*. Food Standards Australia New Zealand, Canberra, Australia

Grayot, J.D. (2019). Dual process theories in behavioral economics and neuroeconomics: a critical review. *Review of Philosophy and Psychology*, *11*, 105-136. https://doi.org/10.1007/s13164-019-00446-9

Hahn, U. & Harris, A. (2014). What does it mean to be biased: Motivated reasoning and rationality. *Psychology of Learning and Motivation*, *61*, 41-102. https://doi.org/10.1016/B978-0-12-800283-4.00001-0

Handley, S. J., & Trippas, D. (2015). Dual processes and the interplay between knowledge and structure: A new parallel processing model. *Psychology of Learning and Motivation, 62,* 33-58. https://doi.org/10.1016/bs.plm.2014.09.002

Hayes, B. K., Dunn, J. C., Joubert, A., & Taylor, R. (2017). Comparing single- and dual-process models of memory development. *Developmental Science, 20,* e12469. https://doi.org/10.1111/desc.12469

Hersey, J.C., Wohlgenant, K.C., Arsenault, J.E., Kosa, K.M., & Muth, M.K. (2013). Effects of front-of-package and shelf nutrition labelling systems on consumers. *Nutrition Reviews*, *71*(1), 1-14. https://doi.org/10.1111/nure.12000

Hleborodova, V. (2018, May 17). Fresh criticism for health star ratings system. *Canstar Blue*. Retrieved from https://www.canstarblue.com.au/news/health-star-ratings-criticised

Keren, G. (2013). A tale of two systems: A scientific advance or a theoretical stone soup? Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, *8*(3), 257-262. https://doi.org/10.1177/1745691613483474

Kruglanski, A.W. & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*(1), 97-109. https://doi.org/10.1037/a0020762

Lai, J.C., Harrison, A., Gao, H. & Becher, S. (2019, June 18). Why the Australasian Health Star Rating needs major changes to make it work. *The Conversation*. https://theconversation.com/why-the-australasian-health-star-rating-needs-major-changes-to-make-it-work-114581

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6,* 312-319. https://doi.org/10.3758/BF03197461

Loftus, G. R., Oberg, M. A., & Dillon, A. M. (2004). Linear theory, dimensional theory, and the face-inversion effect. *Psychological Review, 111*, 835-83. https://doi.org/10.1037/0033-295X.111.4.835

Neal, B., Crino, M., Dunford, E., Gao, A., Greenland, R., Li, N., Ngai, J., Mhurchu, C.N., Pettigrew, S., Sacks, G., Webster, J., & Wu, J.H.Y. (2017). Effects of Different Types of Front-of-Pack Labelling Information on the Healthiness of Food Purchases – A Randomised Controlled Trial. *Nutrients*, *9*(12):1284. https://doi.org/10.3390/nu9121284

Newell, B. R., Dunn, J. C., & Kalish, M. L. (2010). The dimensionality of perceptual category learning: A state-trace analysis. *Memory & Cognition*, *38*, 563-581. https://doi.org/10.3758/MC.38.5.563

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, *11*(6), 988-1010. https://doi.org/10.3758/bf03196730

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in Cognitive Science*, *2*, 511-527. https://doi.org/10.1111/j.1756-8765.2010.01096.x

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80,* 34-72. https://doi.org/10.1016/j.cogpsuch.2015.05.001

Sanjari, S. S., Jahn, S., & Boztug, Y. (2017). Dual-process theory and consumer response to front-of-package nutrition label formats. *Nutrition Reviews, 75*(11), 871-882. https://doi.org/10.1093/nutrit/nux043

Schacter, D. L., & Tulving, E. (1994). *What are the memory systems of 1994?* In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (p. 1-38). The MIT Press

Schneider, G. & Pocheptsova, A. (2020). Should we trust front-of-package labels? How food and brand categorization influence healthiness perception and preference. *Journal of the Association for Consumer Research*, *5*(2). https://doi.org/10.1086/707929

Shepherd, R., Sparks, P., Bellier, S., & Raats, M. M. (1991) The effects of information on sensory ratings and preferences: The importance of attitudes. *Food Quality and Preference, 3*(3), 147-155. https://doi.org/10.1016/0950-3293(91)90051-F

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3-22. https://doi.org/10.1037/0033-2909.119.1.3

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108-131. https://doi.org/10.1207/S15327957PSPR0402_01

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4*(3), 232-243. https://doi.org/10.1162/jocn.1992/4/5/232

Stanovich, K.E., & Toplak, M. E. (2012). Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind & Society*, *11*(1), 3-13. https://doi.org/10.1007/s11299-011-0093-6

Stephens, R.G., Dunn, J.C., Hayes, B.K., & Kalish, M.L. (2020). A test of two processes: The effect of training on deductive and inductive reasoning. *Cognition*, *199*. https://doi.org/10.1016/j.cognition.2020.104223

Stephens, R. G., & Kalish, M. L. (2018). The effect of feedback delay on perceptual category learning and item memory: Further limits of multiple systems. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 44*(9), 1397-1413. https://doi.org/10.1037/xlm0000528

Tanaka, J. W., & Gordon, I. (2011). *Features, configuration, and holistic face processing.* In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.). *The Oxford handbook of face perception* (pp. 177-194). New York, NY: Oxford University Press

The Health Star Rating. (2020, August 12). *Formal review of the system after five years of implementation (June 2014 to June 2019)*. Retrieved from http://www.healthstarrating.gov.au/internet/healthstarrating/publishing.nsf/Content/formal-review-of-the-system-after-five-years

Thompson, V.A., & Johnson, S.C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning*, *20*(2), 215-244. https://doi.org/10.1080/13546783.2013.869763

Thompson, V.A., Prowse Turner, J.A., & Oennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107-140. https://doi.org/10.1016/j.cogpsych.2011.06.001

van Harpen, E., Seiss, E., & van Trijp, H.C.M. (2012). The role of familiarity in front-of-pack label evaluation and use: A comparison between the United Kingdom and The Netherlands. *Food Quality and Preference*, *26*(1), 22-34. https://doi.org/10.1016/j.foodqual.2012.03.003

Wagenmakers, E. J., Krypotos, A. M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition, 40*(2), 145-160. https://doi.org/10.3758/s13421-011-0158-0

World Health Organization. (2019). *World health statistics 2019: Monitoring health for the SDGs, sustainable development goals*. Geneva, Switzerland: Author

Appendix A

Experiment Instructions: Explanation of the Health Star Rating and Nutrition Information Panel

Health Star Rating

The Health Star Rating is a front-of-pack labelling system that rates the overall nutritional profile of packaged food and assigns it a rating from 0.5 star to 5 stars. It provides a quick, easy, standard way to compare similar packaged foods. The more stars, the healthier the choice.

NOTE: The rating is based on most of the components included in the Nutrition Information Panel, but is also based on other factors such as the amount of fruit and vegetable content, and the directions on the label for preparing the food. Therefore, when assessing the healthiness of foods, it is important to also consider information beyond the Health Star Rating, such as the Nutrition Information Panel.

Example of Health Star Rating:

Nutrition Information Panel

Nutrition Information Panels provide information on the average amount of energy, protein, fat, saturated fat, carbohydrate, sugars and sodium in the food.

Here is an example Nutrition Information Panel, with a Guide to interpreting the different values:

Nutrition Information

Servings per package – 6
Serving size – 90g

	Per serve	Per 100g
Energy	**1297kJ**	**1441kJ**
Protein	8.4g	9.3g
Fat		
Total	1.0g	1.1g
Saturated	0.3g	0.3g
Carbohydrate		
Total	44.6g	49.6g
Sugars	10.6g	11.8g
Sodium	194mg	215mg

If comparing nutrients in similar food products, use the per 100g column.

Energy > 1340kJ is considered high

Saturated Fat > 3.0g is considered high

Sugars > 18g is considered high

Sodium (Salt) > 450mg is considered high

Generally, healthier foods are lower in energy, saturated fat, sugars and sodium (salt). These aspects of food are associated with increasing the risk factors for chronic diseases.

We will show you a summary of this guide during the experiment task, for reference.

Appendix B

Mean Response Time Measured for Each Condition

Table B1

Response Time for Response 1

HSR	Logo	NIP	N	*M*	*SD*
Healthy	Healthy	Healthy	60	8.1	11.0
Healthy	Healthy	Unhealthy	60	10.9	25.1
Healthy	Unhealthy	Healthy	60	10.8	23.4
Healthy	Unhealthy	Unhealthy	60	8.8	11.8
Unhealthy	Healthy	Healthy	60	7.4	6.6
Unhealthy	Healthy	Unhealthy	60	9.4	16.9
Unhealthy	Unhealthy	Healthy	60	7.3	6.4
Unhealthy	Unhealthy	Unhealthy	60	6.3	6.1

Table B2

Response Time for Response 2

HSR	Logo	NIP	N	*M*	*SD*
Healthy	Healthy	Healthy	60	9.9	5.4
Healthy	Healthy	Unhealthy	60	11.9	15.0
Healthy	Unhealthy	Healthy	60	12.4	13.0
Healthy	Unhealthy	Unhealthy	60	10.1	7.4
Unhealthy	Healthy	Healthy	60	12.6	10.3
Unhealthy	Healthy	Unhealthy	60	10.8	7.6
Unhealthy	Unhealthy	Healthy	60	11.2	12.7
Unhealthy	Unhealthy	Unhealthy	60	10.3	10.5

Appendix C

Four-way Repeated Measures ANOVA Results

Effect	**DFn**	**DFd**	***F***	***p***	***ges***
Rating Type	1	59	21.802	< 0.001*	0.019
HSR	1	59	179.263	< 0.001*	0.374
Logo	1	59	72.616	< 0.001*	0.058
NIP	1	59	155.551	< 0.001*	0.190
Rating Type x HSR	1	59	83.452	<0.001*	0.163
Rating Type x Logo	1	59	37.649	<0.001*	0.028
HSR x Logo	1	59	0.080	0.778	0.000
Rating Type x NIP	1	59	155.001	<0.001*	0.113
HSR x NIP	1	59	5.864	0.019*	0.002
Logo x NIP	1	59	0.694	0.408	0.000
Rating Type x HSR x Logo	1	59	5.870	0.018*	0.001
Rating Type x HSR x NIP	1	59	0.317	0.575	0.000
Rating Type x Logo x NIP	1	59	0.118	0.732	0.000
HSR x Logo x NIP	1	59	1.753	0.191	0.000
Rating Type x HSR x Logo x NIP	1	59	0.071	0.791	0.0000096

Note. * = $p < 0.05$; *ges* = “generalised eta squared”.