Integrating dual-process and pragmatic theories for the processing of verbal and numerical food quantifiers

Dawn Liu

A thesis submitted for the Degree of Doctor of Philosophy in Psychology

Department of Psychology

University of Essex

9 January 2020

Declaration

I declare that this thesis, '*Integrating dual-process and pragmatic theories for the processing of verbal and numerical food quantifiers*', and the work presented in it is my own and has been generated by me as the result of my own original research. All sources of information referred to in the thesis have been acknowledged. None of the work in this thesis has been submitted for a higher degree at this or any other University or institution.

Submitted by: Dawn Liu

Signature:

Acknowledgements

Ironically, after producing over 50,000 words about quantifiers, I find it virtually impossible to quantify my gratitude for the wealth of knowledge exchange, mentorship, and inspiration I encountered while doing this work. These acknowledgements are by no means complete, but they are necessary.

Above all, I have immense gratitude for my supervisor, Dr Marie Juanchich. All the words in this thesis are insufficient to express what a pillar of support, guidance, and encouragement you have been. I would need a whole book (and then some) to extol your virtues. For now, I shall confine myself to two words and hope you can draw the relevant pragmatic inference about the depth of my gratitude: thank you!

I owe great thanks as well to my second and third supervisors, Dr Miroslav Sirota and Professor Sheina Orbell: I have benefitted greatly from your excellent supervision, expert advice, and pastoral support over the last three years. Thank you for encouraging me to keep going when things got tough! I thank also Professor Linda Moxey and Dr Jonathan Rolison for kindly examining my thesis, as well as Professor Michael Steinke for acting as independent chair.

To my wonderful colleagues in the Essex psychology department —and there are too many of you to name individually —thank you for the tea and cakes, for sharing your skills, for listening, for researcher solidarity, and daily reminders that this career is the right one for me. To friends and family, thank you for enthusiastic cheerleading from 7,000 (or 'many'!) miles away, despite not quite understanding what I do all day, and for believing in me all the same.

And to my dearest Angus, for all the big and small ways you have made a difference to my life and my PhD: thank you for appreciating and indulging my quirky researcher habits that never 'switch off'; for advice on producing, submitting, and life after the PhD; and for being the kind of academic I aspire to be.

General Abstract

This thesis investigated how using a verbal or numerical quantifier format affects people's psychological judgement and decision-making in three areas: interpretations, attention to, and evaluation of quantified information. These formats are commonly used in nutrition communication, but there is a paucity of evidence in the literature on how they each affect judgement and decision-making processes, and how this could affect current practices in applied communications. Over 14 pre-registered studies, this research drew on two previously independent theoretical frameworks, dual-process theory and pragmatic theory, to explain the differences in people's processing of verbal and numerical quantifiers, with a specific focus on quantifiers used to convey food information. Chapter 1 gives a brief overview of the theoretical and applied literature. Chapter 2 (Experiments 1-2) discusses the substantial inter-individual variance found in interpretation of verbal nutrition quantifiers, and a general misalignment between consumer interpretations and standard guidelines. Chapters 3-4 (Experiments 3-6) show using multiple measures of processing (response time, decision performance, subjective effort, reliance on contextual information, and performance under cognitive load) that verbal quantifiers are not necessarily more intuitively processed than numerical quantifiers, but people may make better decisions quicker with numerical quantifiers. Chapter 5 (Experiment 7) shows how people's attention to quantified attributes is greater with verbal than numerical quantifiers. Finally, Chapters 6-7 (Experiments 8-14) leverage the attribute framing effect, where positive frames (e.g., 'energy') are preferred to negative ones (e.g., 'calories'), to compare dual-process and pragmatic theories of quantifier processing. Overall, verbal quantifiers can increase the influence of informational context on a person's decision, however the process involves both affect-driven responses and the extraction of implicit information from the communicative choices of a speaker. These findings can be applied to better frame quantifiers presented on nutrition labels such that consumers receive accurate and useful information.

Author's Note

The empirical chapters of this thesis (2 through 7) were written as independent manuscripts. This facilitated submission of the work to peer-reviewed journals for publication. Although each reports on independent studies, some content in the introductions of these chapter may overlap with each other and the General Introduction (Chapter 1) and General Discussion (Chapter 8) of the thesis. However, the references for each chapter are summarised only at the end of the thesis.

At time of submission, a revised version of Chapter 2 is published in *Food Quality and Preference*. Chapter 3 is in revision at *Thinking and Reasoning*. Chapter 4 has been accepted for publication by the *Quarterly Journal of Experimental Psychology*. Chapter 5 is being revised for submission to the *Journal of Memory and Language*. Chapter 6 is under review at *Acta Psychologica*, and Chapter 7 is under review at the *Journal of Behavioral Decision Making*.

Contents

List of Figures

List of Tables

Integrating dual-process and pragmatic theories for the processing of verbal and numerical food quantifiers

"Words and numbers are of equal value, for, in the cloak of knowledge, one is warp and the other woof. It is no more important to count the sands than it is to name the stars."

- Norton Juster, *The Phantom Tollbooth*

Chapter 1: General Introduction

1.1 The Communication of Quantities

One rarely gets through a day without relying on quantifiers. At breakfast, one might consider the nutrients in the meal, which relies on a quantifier to describe for example, the amount of fat, be it a precise numerical amount (e.g., '5% fat'), or a verbal quantifier (e.g., 'low fat'). One might go to the doctor's office and be warned about the magnitude of a risk from taking medication using a frequency quantifier such as 'common' (or 10%; Berry et al., 2003). One could look up the weather forecast for the next day and find the uncertainty surrounding a predicted rain forecast expressed with a probabilistic quantifier, such as 'likely' (or 60%; Knox, 1969; Sink, 1995). In this way, quantifiers serve an important function in daily communication (Moxey & Sanford, 2000), by scoping information (Kiss & Pafel, 2017). Therefore, understanding quantifiers and what they convey is crucial to making informed decisions (Juanchich et al., 2012; Mandel, 2015). This thesis is motivated by this need to better understand how people process proportional quantifiers of different formats, specifically numerical percentages, and their relevant verbal (also known as 'linguistic') expressions. I examine how using verbal or numerical quantifiers in food information affects interpretation, attention, and evaluation processes for food.

1.1.1 Theoretical underpinnings of verbal quantifier research

Although quantification is most commonly understood to involve numerical calculations (Moxey & Sanford, 1992), the vast majority of quantified communication in daily life takes place using verbal quantifiers (Juanchich et al., 2019; Moxey & Sanford, 2000). Indeed, people often prefer using verbal quantifiers over numerical ones in natural language (Wallsten et al., 1993). Since 1970, a concerted line of research that focused on the interpretation of verbal probabilities and the subsequent impact on decision-making highlights an increasing acknowledgement of the importance of verbal quantifiers in communication (e.g., Budescu

& Wallsten, 1995; Druzdzel, 1989; Juanchich et al., 2019; Teigen & Brun, 2003).

Within the psychological literature, research on quantifiers is largely divided into two strands. The first takes a judgement and decision-making (JDM) approach, and concerns whether people are 'rational' (i.e., accurate) in interpreting verbal quantifiers. Here, research across various domains shows that verbal quantifiers are 'vague', with participants providing translations from verbal-numerical quantifiers that vary widely between individuals and across contexts, and tend to misalign with intended translations (e.g., Berry, 2006; Beyth-Marom, 1982; Budescu et al., 2014; Budescu & Wallsten, 1985; MacLeod & Pietravalle, 2017; Webster et al., 2017; see Druzdzel, 1989 for a review). The other strand takes a psycholinguistics approach, and focuses on how people derive functional linguistic interpretations of verbal quantifiers (e.g., by making focal inferences; Moxey & Sanford, 2000). Here, a key finding is that verbal quantifiers are directional: participants tend to explain 'it is likely to rain' (positive focus) with reasons why it would rain, whereas 'it is unlikely to rain' (negative focus) evokes reasons why it would not rain (Sanford et al., 2007; Teigen & Brun, 1995).

The two different approaches to research on verbal quantifier have developed different accounts that lead to different perspectives on how people evaluate verbal quantifiers compared to numerical ones. The JDM approach predicts that people are less precise and more biased with verbal quantifiers (Windschitl & Wells, 1996). In contrast, the psycholinguistics approach predicts that people reach reasonable decisions with verbal quantifiers because they factor in implicit information from the framing of the verbal quantifier (Teigen & Brun, 1999). Some research has begun to compare verbal quantifier interpretations from the JDM and psycholinguistics perspectives, finding that people's numerical interpretation of a verbal quantifier can predict whether they will use a verbal quantifier of positive or negative focus (Budescu et al., 2003). People also use contextual factors such as the base rate of the quantified event to determine which verbal quantifier focus to use (Juanchich et al., 2013). This type of comparative research integrates previously separate theoretical accounts. The next step to further the research is to apply each theory to the psychological processes involved in a decision with quantifiers. Each theory provides a different lens by which to view the decision-making process. Integrating theories thus provides a more holistic approach to mapping the psychological processing of quantifiers in decision-making (Gigerenzer, 2011; Glöckner & Betsch, 2011).

1.1.2 Applying research in quantifier processing to practice

A further gap in the literature on verbal and numerical quantifiers involves connecting methods and outcomes from cognitive and behavioural research to practical applications involving quantity communication. How people process quantifiers, in terms of interpretation, attention, and evaluation, is crucial for their decision-making. For example, to decide whether a food is healthy, one would need to interpret a ‘low fat’ label, pay attention to the different components of the information (e.g., focusing on ‘low’ or focusing on the ‘fat’), and use that information to evaluate the food. Whether the label uses a verbal (‘low’) or numerical (‘5%’) label could affect each of these processes. However, applied research in nutrition labelling lacks evidence on how the quantifier format affects processing. Most applied research compares labels in their existing real-world formats (e.g., Borgmeier & Westenhoefer, 2009; Emrich et al., 2014; Savoie et al., 2013; van Herpen et al., 2014; see also Byrd-Bredbenner et al., 2000; Campos et al., 2011; Cowburn & Stockley, 2005 for reviews). These labels vary quantifier format in an unsystematic manner, limiting the conclusions that can be drawn about the effect of quantifier format. Thus there is a need for more systematic investigations that derive and test specific, theoretically-driven hypotheses within the applied context.

1.1.3 Project goals

The overarching goal of this project was to address the theoretical and applied gaps in current research on decision-making processes with quantifiers. Specifically, I investigated proportional quantifiers (numerical percentages and their corresponding verbal quantifier expressions; Keenan & Paperno, 2012) in terms of three key psychological processes: first, the interpretation of quantifiers, defined as the meaning that people assign to a quantifier; second, attention to information, defined as the amount and focus of processing given to the quantified information; and third, evaluation, defined as the final judgement or decision reached. I drew from two influential theories in JDM and psycholinguistics as a framework to derive specific hypotheses for these processes: dual-process theory (De Neys, 2017b; Evans, 2008) and pragmatic theory (Horn & Ward, 2006). Finally, I contextualised my investigation by applying it to the practical problem of communicating nutritional information effectively.

This chapter is structured as follows. First, I will review past work on the interpretation of verbal and numerical quantifiers. Second, I will give an overview of the literature on dual-process theory and pragmatic theory, which will form the theoretical framework to explain differences in the way people attend to and evaluate verbal and numerical quantifier formats. Third, I will outline some of the practical applications of studying verbal vs. numerical quantifiers in nutrition communication. Finally, I will present the structure of this thesis, along with a brief explanation of the goals of the studies conducted in each chapter. I end the chapter with a brief discussion of the commitment of this work to open science practices.

1.2 Interpretations of Verbal and Numerical Quantifiers: How High is 'High'?

Researchers have attempted for decades to map verbal quantifiers onto numerical ones (e.g., Borges & Sawyers, 1974; Budescu et al., 2003; Budescu & Wallsten, 1985; Juanchich et al., 2013; Renooij & Witteman, 1999). This research shows that there is not a one-to-one translation between the two quantifier formats (Budescu & Wallsten, 1995; Juanchich et al., 2019; Reagan et al., 1989). Verbal quantifiers can be measured using a range of numerical equivalents (known as the 'membership function'; Budescu & Wallsten, 1985), the values of which differ among individuals and across contexts (Berry et al., 2002; Budescu et al., 2014; Budescu & Wallsten, 1985; MacLeod & Pietravalle, 2017; Teigen & Brun, 1999). For example, a 'high' chance was translated on average as 83% across a participant sample, but actual translations per individual could vary from 65-95% (MacLeod & Pietravalle, 2017). Further, a high chance to contract a disease might mean 83%, but a high chance of snow could mean 95% (Patt & Schrag, 2003). Interestingly, even when presented with translation guidelines for what numerical values should be interpreted from a verbal quantifier, people still produce numerical translations that are inconsistent with the given guidelines (Budescu et al., 2014). People appear to have automatic preferences for how they initially interpret linguistic information (Altmann, 1998).

It is generally agreed that verbal quantifiers are vague and numerical quantifiers precise (Budescu & Wallsten, 1995), although people do display some variability in their perceptions of numerical quantifiers (Budescu et al., 1988). Yet the interpretation of numerical quantifiers can also be subject to changes in the context. For example, people may see the same objective information (e.g., a visual display of the frequency of 1,000 apple trees, of which a number produce bad apples) and give higher numerical estimates of bad trees when the bad apples are poisonous than when they are simply sour (Harris et al., 2009). Because the research on verbal-numerical quantifier translations has typically been unidirectional, translating from verbal quantifiers to numerical, less is known about how numerical quantifiers might be mapped onto verbal ones, or if indeed people do interpret numerical quantifiers as precise amounts. Chapter 2 of this thesis will address these questions. In the next section, I will also describe the methodological challenges caused by interpretational vagueness, and in the subsequent chapters, introduce methodological adjustments to control for comparative vagueness of verbal and numerical quantifiers in empirical work.

1.3 Decision-Making with Verbal and Numerical Quantifiers: How Do People Process the Information?

Behavioural evidence indicates that there are differences in how people reach decisions with verbal and numerical quantifiers, though the exact nature of these differences varies across studies (e.g., Budescu & Wallsten, 1990; Childers & Viswanathan, 2000; González-Vallejo et al., 1994; Moxey & Sanford, 2000; Viswanathan & Childers, 1996). Compared to numerical quantifiers, verbal quantifiers are more vague (Budescu & Wallsten, 1995; Sen, 1998; Zimmer, 1983), but easier to contextualise (Erev & Cohen, 1990; Viswanathan & Childers, 1996;

Viswanathan & Narayanan, 1994). Interestingly, although people tend to prefer communicating in verbal formats, the format in which they prefer to receive quantified information varies over contexts and communicative purposes (Erev & Cohen, 1990; Olson & Budescu, 1997; Wallsten et al., 1993). Quantifier format consistently affects judgement and decision-making outcomes, but the direction of the effect is inconsistent. Despite showing different decision patterns, neither format shows a clear-cut advantage over the other in terms of aggregated decision quality (Budescu et al., 1988; González-Vallejo et al., 1994). It is therefore necessary to look beyond behavioural outcomes to consider the more nuanced cognitive processes that may take place to produce the decision output (Payne & Bettman, 2007). This would illuminate the underlying explanations for research findings to date and allow better predictions for the conditions where a verbal or a numerical quantifier would produce better judgements and decisions.

This section will review the literature that compares verbal and numerical quantifiers, with a focus on attention and evaluation of information with different quantifier formats. I approach this discussion from two contrasting perspectives: a dual-process theory perspective, and a pragmatic theory perspective. Each of these theories encompasses many sub-theories that explain different aspects and levels of information processes (De Neys, 2017b; Horn & Ward, 2006). Dual-process theories have a long tradition in socio-cognitive psychology and JDM research and tackle information processing at the basic cognitive, or mental, level (Gilovich & Griffin, 2010). In this thesis, I focus on dual-process theory's basic conception of two distinct processing systems (intuition and analysis; Evans, 2008). This perspective fits many reported differences in verbal and numerical quantifiers (Windschitl & Wells, 1996). Pragmatic theories draw from psycholinguistics and span an overarching goal of comparing the semantic vs. practical meaning people derive from verbal quantifiers (Horn, 2006). In this thesis, I focus within pragmatic theory on the domain of implicature, which highlights that a communicator conveys more than the logical meaning of their information content (Horn, 2006). This perspective acknowledges that a selected quantifier is a language choice, which must therefore be understood in that context (Sher & McKenzie, 2006).

1.3.1 Intuition vs. analysis: quick and biased or slow and rational?

The brain does not appear to process information in only one manner. At times, processing appears to be quick and effortless, for example, for most adults, it is easy to read a word on a food label. At other times, information requires more effort to process, such as trying to calculate the average of five numbers. The idea that two styles of processing exist has permeated JDM research for many decades (Gilovich & Griffin, 2010). In its most basic form, dual-process theory proposes that the two styles differ in terms of consciousness, automaticity, and the amount of cognitive effort involved (Evans, 2008). The 'intuitive' process (also termed 'heuristic'; Zuckerman & Chaiken, 1998, or 'experiential'; Epstein et al., 1996) is unconscious, automatic, and quick, often driven by affective cognitions about the decision, such as feelings that it is 'right'. In contrast, the 'analytical' process (also termed 'deliberative'; Plessner et al., 2008, 'rational'; Epstein et al., 1996, or 'systematic'; Zuckerman & Chaiken, 1998) reflects conscious, effortful processing that operates slower and is driven by rational cognitions about the decision, such as deliberate reasoning (see Evans, 2008 for a discussion of these concepts, but also Glöckner & Witteman, 2010b for alternative viewpoints, and Melnikoff & Bargh, 2018 for a critique)[1].

How verbal and numerical quantifiers fit intuitive and analytical processes. Three reasons suggest dual-process theory as an explanation for differences in decisions with verbal and numerical quantifiers. First, the way people use verbal and numerical quantifiers seems to fit a dual-process theory classification in terms of the level of automaticity and consciousness involved. Words are usually processed automatically and unconsciously, as seen in tasks like the Stroop effect: the conflict people face in naming, for example, the colour of the word 'blue'

[1]The terms used to distinguish the two processing styles are not consistent across the literature. For consistency, and to avoid confusion in cases where the other labels conflict with alternative definitions (e.g., 'heuristics' are also used to describe automatic mental shortcuts to avoid more costly cognition Kahneman, 2011, or intentional strategies to simplify a more complex calculation Gigerenzer, 2016; these two definitions are not interchangeable or necessarily equivalent to intuition), I have chosen to discuss the styles in terms of 'intuition' and 'analysis' to reflect their relation to the mental processes involved.

presented in yellow font, is caused by their automatic generation of the meaning of 'blue' (MacLeod, 1991). In contrast, people often need to expend deliberate effort to process numerical information such as performing calculations, where each number must be held in working memory during the computation (DeStefano & LeFevre, 2004). While people are able to derive an automatic sense of where a verbal quantifier such as 'high' lies on an evaluative scale, thus make such evaluations quickly (Viswanathan & Childers, 1996) and remember them better (Scammon, 1977), they have difficulty processing numbers when they need to be understood in an evaluative context —for instance, identifying whether 60% is a good or bad quantity —and do so slower and with poorer memory for the evaluation (Scammon, 1977; Viswanathan & Childers, 1997). Studies investigating deliberation over quantified information also find that people consciously mention numerical quantifiers in their deliberations, but show signs of unconsciously using verbal quantifiers, for example mentioning that food labels with verbal quantifiers were easier to use in their decision-making (Malam et al., 2009), or mentioning information that was only provided in verbal format when they explained their decisions *after* their main deliberation (Ang & Trotman, 2015). Based on these examples, verbal quantifiers fit the automatic and unconscious nature of intuitive processes, while numerical quantifiers fit the controlled and deliberate nature of analytical processes.

The second reason dual-process theory could explain verbal and numerical quantifier processing is that people prefer to make decisions with verbal than numerical quantifiers when the task suits intuitive rather than analytical processing (Wallsten et al., 1993). For example, intuitive tasks are subjective (e.g., judging facial expressions; Ayal et al., 2015; Rusou et al., 2013), vague (e.g., containing uncertain or imprecise information; Hammond, 1988), and/or rely on affective judgements (e.g., personal preferences; Wilson & Schooler, 1991). There is evidence that people prefer verbal quantifiers in tasks that have these characteristics (Nicolas et al., 2010; Wallsten et al., 1993), suggesting that verbal quantifiers cue an intuitive style compatible with the task. Conversely, people prefer numerical quantifiers when the task involves objective, precise, and analytical reasoning (Budescu & Wallsten, 1990). Numerical quantifiers may thus cue a compatible analytical style. Dunwoody et al. (2000) concluded that people process numerical information analytically based on a number of factors including slower decisions, higher consistency of decisions, and better ability to explain decisions. However, they compared the numerical information to pictorial information, thus they did not have a direct test between a task with numerical and verbal quantifiers. A study that did compare the format of gamble probabilities found that people tended to pick gambles that lead to more positive outcome values (i.e., higher payoffs) when given verbal probabilities, but gambles that matched the probability values (i.e., higher success likelihoods) when given numerical probabilities (González-Vallejo et al., 1994). These findings support the view that verbal quantifiers suit a more affect-driven, intuitive processing style, while numerical quantifiers suit a more reason-based, analytical processing style.

The third reason why dual-process theory could explain verbal and numerical quantifier processing is that in comparison to numerical quantifiers, verbal quantifiers appear to elicit a more associative processing approach that is prone to common judgement biases (Welkenhuysen et al., 2001; Windschitl & Wells, 1996)[2]. A dual-process explanation for such judgement biases proposes that people are more influenced by affective information when relying on intuition (Levin & Gaeth, 1988; Slovic et al., 2007). For example, in the 'framing effect', where people judge the same information differently depending on whether it is framed in a positive or negative manner, people are posited to encode affective information about an attribute's frame, which later primes them to judge a positively-framed target more favourably, and vice versa (Levin & Gaeth, 1988). Using analytical

[2]Although the view that intuition produces flawed cognitions (Tversky & Kahneman, 1974) is becoming increasingly challenged (Bago & De Neys, 2019; Gigerenzer, 2016; Plessner & Czenna, 2008), the challenge to the original premise of intuitive biases being flawed is that the biases are erroneous by rational, but not ecological standards (Gigerenzer & Todd, 1999). Therefore, a tendency for verbal quantifiers to produce more 'biased' judgements than numerical quantifiers still fits an intuitive/analytical distinction. Whether such judgements are of poorer quality depends on the specific context, although stricter versions of dual-process reasoning (e.g., taking weighted analysis as the gold standard for decision performance Czerlinski et al., 1999) might posit that intuitive verbal quantifiers produce worse decision performance than analytical numerical ones.

processing should attenuate the bias (Thomas & Millar, 2012), and this is indeed seen with numerical vs. verbal quantifiers. Welkenhuysen et al. (2001) found that participants were more likely to get a medical test when told they had a moderate chance of having a baby with cystic fibrosis (negative frame) than a high chance of having one without (positive frame), but they did not show this effect with the corresponding numerical quantifiers (25% and 75%). People are also less susceptible to the denominator neglect bias, where a ratio is judged on the basis of its numerator without considering the fraction in its entirety, when uncertainty is measured using numerical than verbal quantifiers: a ‘1 in 10’ chance is ‘very unlikely’ whereas a ‘10 in 100’ chance is merely ‘unlikely’, but both fractions are described as 10% likelihoods (Windschitl & Wells, 1996). The greater susceptibility to judgement biases with verbal quantifiers therefore suggest that this format encourages more reliance on affective information in decision-making, in line with an intuitive processing style.

Challenges for a dual-process theory of quantifier processing. The three reasons given above suggest dual-process theory as a candidate framework to explain JDM differences for verbal and numerical quantifiers. However, certain findings challenge the duality of the quantifier formats. Based on the traditional correlates of dual-process theory, intuitive processes should operate faster, require less subjective effort, and result in poorer performance than analytical processes (Evans, 2008). Although these views are being challenged (Melnikoff & Bargh, 2018), they remain core hypotheses of the theory. Thus, if verbal quantifiers and numerical quantifiers are intuitive and analytical, verbal quantifiers should be processed more quickly, with less effort, and result in poorer performance than numerical quantifiers.

A challenge for a dual-process theory perspective of verbal and numerical quantifier processing is that direct comparisons of verbal and numerical quantifiers provide mixed evidence on the speed, effort, and accuracy of decision-making. Studies have found quicker decision times for verbal than numerical quantifiers, but also vice versa. People choose gambles with verbal quantifiers (Budescu & Wallsten, 1990) and evaluate verbal quantifiers (Viswanathan & Childers, 1996) faster than numerical quantifiers. However, people are quicker to compare the magnitude of two numerical than verbal quantifiers (Jaffe-Katz et al., 1989; Viswanathan & Narayanan, 1994), and spend more time viewing verbal than numerical quantifiers (Viswanathan & Childers, 1996). Yet other studies report no reaction time differences in picking verbal or numerical gambles (González-Vallejo et al., 1994), or that the speed of the response depends on other factors. For instance, people choose quicker between items with verbally quantified attributes when the items are distinct, but choose quicker with numerical quantifiers when the items are similar (Stone & Schkade, 1991). In terms of effort, people report that they find verbal quantifiers less effortful to process than numerical quantifiers (Cowburn & Stockley, 2005; Peters et al., 2009). However, information recipients prefer numerical over verbal quantifiers, but information communicators prefer the reverse (Erev & Cohen, 1990; Wallsten et al., 1993; Xu et al., 2008). Because preference for information is often an indicator of the level of processing difficulty (Dunn et al., 2016; Shenhav et al., 2017), the 'communication mode preference' (CMP; Erev & Cohen, 1990) suggests that whether verbal or numerical quantifiers are more effortful depends on whether one is giving or receiving the information. The evidence for whether people perform better with verbal or numerical quantifiers is similarly mixed, with studies finding that numerically-quantified product attributes are recalled more accurately (Viswanathan & Childers, 1996), other studies finding people can remember nutritional information more accurately when given verbal quantifiers (Scammon, 1977), and still others reporting that the payoffs from participants' selected gambles were not on average greater with verbal or numerical quantifiers (González-Vallejo et al., 1994).

The mixed evidence regarding how verbal and numerical quantifiers differ on response time, effort, and decision performance may reflect certain methodological issues in comparing the two formats along individual measures of processing style. First, the studies reviewed employed different methods to elicit their dependent measures. For example, evaluating quantifiers in a practical context and deciding on a gamble (Budescu & Wallsten, 1990; Viswanathan & Childers, 1996) involve more complex processes than comparing the magnitude of two quantifiers (Viswanathan & Childers, 1996) or simply reading the quantifiers (as could be the case in Viswanathan & Childers, 1996's study). A direct comparison of decision processes for verbal and numerical quantifiers is thus needed, which outlines the goal of the decision task and what should be expected if an intuitive or analytical approach to the task is taken.

Second, previous study methods have not fully accounted for the vagueness of verbal quantifiers. Although it is established that the meaning of verbal quantifiers vary greatly between individuals, past research has compared the quantifier formats based on the average numerical translation of the verbal quantifiers (e.g., Jaffe-Katz et al., 1989; Viswanathan & Childers, 1996), or interpretations selected by the researcher (e.g., Stone & Schkade, 1991). In an attempt to address individual differences in quantifier interpretation, (Budescu & Wallsten, 1990) paired participants such that each decision-maker received quantifiers from another participant, however they found that the recipients tended to interpret the verbal quantifiers differently from the communicator. One might expect vague interpretations of verbal quantifiers to result in a poorer performance (e.g., lower gamble payoffs) with verbal than numerical quantifiers, due to the lower precision. However, this was not the case in Budescu & Wallsten (1990)'s study, suggesting that over- or underestimations due to verbal vagueness tend to average out over a sample. Nonetheless, in studies that measured magnitude comparisons of verbal or numerical quantifiers, verbal vagueness might explain why people perform less well when the verbal quantities are less distinct. To rule out verbal vagueness as an explanation for processing differences, it is thus necessary to use methods that should only produce differences between verbal and numerical quantifiers if people use a different processing style.

Third, the body of research so far has tended to look at each measure in isolation, and not as a direct test of the hypotheses of dual-process theory. Response times, effort, and performance are correlates of dual-process theory, rather than the defining characteristic of intuitive and analytical processes (Evans & Stanovich, 2013). Although dual-process theory predicts that intuitive processes should be quicker, less effortful, and less accurate, these features do not isolate an intuitive process (Pennycook, 2017). Rather, a more critical test of dual-process theory should assess the demands of the process on working memory (Evans & Stanovich, 2013). Intuitive processes should be autonomous and create minimal demands on working memory, while analytical processes should involve cognitive decoupling (e.g., hypothetical thinking) and thus create a higher demand on working memory (Evans & Stanovich, 2013). Adding additional memory load should therefore suppress analytical processing ability (Białek & De Neys, 2017; De Neys, 2006; Trémolière et al., 2014). Some evidence suggests that memory load could affect numerical processing more than verbal: people are less accurate at remembering numerical than verbal quantifiers when the number of quantities involved increases (Scammon, 1977). However, the standard manipulation of presenting a concurrent cognitive load to impair analytical processing (De Neys, 2006), has not, to my knowledge, been applied to test if the processing of a verbal quantifier would proceed unimpaired under load, while the processing of a numerical quantifier would be impaired by the concurrent load. The current state of research therefore lacks a direct test of dual-process theory that also rules out verbal vagueness as an explanation for processing differences between verbal and numerical quantifiers. I seek to address the research gap, summarised by the three methodological issues outlined above, that could account for the current mixed evidence on measures of quantifier processing in Chapters 3 and 4 of this thesis.

1.3.2 Meaning more than we say: Pragmatic inferences about quantifiers

From a language processing perspective, both words and numbers are symbolic mental representations of concepts (Paivio, 1990). The interpretation of quantifiers, whether verbal or numerical, involves a process whereby meaning is extracted from what is perceived. A key question from a psycholinguistics perspective is what meaning is derived, as well as how much information is encoded in the mental representation. The field of psycholinguistics approaches these questions from two angles: language production (how and why information is communicated in a certain way) and language comprehension (how this information is then understood; Kess, 1992). While dual-process theory is primarily concerned about people's cognitive approach to information processing, pragmatic theory focuses on a different level of processing: how people use linguistic information to fulfil a communicative purpose. This approach acknowledges that language enables communicators to convey not only their informational content, but also what they think their recipient should do with the information (Kess, 1992). In natural communicative circumstances, recipients must, and will, constantly infer details that were not explicitly stated by the communicator, such as a prescription for them to act (Johnson-Laird et al., 1986). These implicatures, drawn based on shared common knowledge and implicit conventions in discourse, allow communication to be delivered in a practically concise fashion (Grice, 1975). For example, if told '*there is little risk associated with taking this medicine*', one could infer that they should just take the medicine, even if this counsel is not explicitly stated. Yet one could also expect a listener to infer from '*a slight risk associated with taking this medicine*' that they should be cautious about taking it. Although the descriptive content of the two statements give a similar level of risk, the recommendations for action are in the opposite direction (Sanford & Moxey, 2003; Schmeltzer & Hilton, 2014; van Buiten & Keren, 2009). By delivering both a warning as well as the magnitude of the likelihood of the risk, verbal quantifiers may perform the dual function of communicating speaker intent over and above informational content (Sanford & Moxey, 2003; van Buiten & Keren, 2009). In contrast, an equivalent numerical quantifier such as 'a 2% risk' could provide greater precision, but be less informative in terms of the speaker's recommendation for action (Keren, 2007). The study of implicature in pragmatic theory thus seeks to address whether information recipients do in fact make these inferences and rely on them in their decision-making.

Pragmatic inference: a process of deriving different information from verbal and numerical quantifiers. Pragmatic implicature offers a different perspective to the processes of attention and evaluation of quantified information. It is not about an individual's cognitive approach to processing the information (i.e., processing style in the dual-process dichotomy) but the embedded meaning that can be extracted while processing the information, and the resulting effect on judgements and decisions (Sanford et al., 2007; Teigen & Brun, 1999). While research on the interpretation of verbal and numerical quantifiers suggests that verbal quantifiers are less informative due to their greater vagueness and variability (Budescu & Wallsten, 1995), pragmatic theory identifies that verbal quantifiers can be more informative than numerical ones by conveying directional information, or 'focus' (Honda & Yamagishi, 2017; Moxey & Sanford, 1986; Teigen & Brun, 1999)[3]. A verbal quantifier can focus attention on different attributes in a statement. For example, '*a few apples were good*' leads people to follow-up with sentences that refer to the good apples (the set of attributes referred to in the statement, or 'positive focus'), whereas people could follow-up on '*few apples were good*' with sentences that refer to the bad apples (the set of attributes not referred to in the statement, or 'negative focus'; Moxey & Sanford, 1986). In contrast, the corresponding numerical quantifier, e.g., '*20% of apples were good*', might be more ambiguous in terms of whether people focus on the good or bad apples. In general, work on the directional focus of verbal probabilities has found that compared to the equivalent verbal probability, numerical probabilities tended to be ambiguous, but typically focused on the referenced attribute: with a 30% chance of success, people might focus on either success or failure, but more on success (Teigen & Brun, 2000). Notably, the equivalent (negative) verbal probability, 'quite uncertain', would direct focus to the chance of failure: people are more likely to give reasons for why a quite uncertain success would *not* happen (Teigen & Brun, 1995).

The ability of verbal quantifiers to direct focus in different directions allows people to make inferences about the quantities and about the communicator's intentions (Honda & Yamagishi, 2017; Keren, 2007; Sanford & Moxey, 2003). In this way, verbal quantifiers can constitute a natural 'frame'. As reviewed in the earlier section on dual-process theories, people appear to be more susceptible to the framing effect with verbal quantifiers (Welkenhuysen et al., 2001; Windschitl & Wells, 1996). According to pragmatic theory, people would still evaluate a positive verbal quantifier (e.g., '*a few people came to the party*') as better than

[3]Two approaches to this research on the implied meaning of quantifiers use different terminology. In their work on uncertainty quantifiers (i.e., probabilities), Teigen & Brun (1995) refer to this feature as '*directionality*'. Moxey & Sanford (1986)'s work (primarily on existential quantifiers, e.g., '*a few*', '*some*') refers to it as '*directional focus*'. There are subtle differences in the way these concepts are implemented, however to avoid confusion, I have opted to discuss this literature in terms of directional focus, because this thesis investigates non-probabilistic quantifiers.

a negative one (e.g., '*few people came to the party*'; Moxey, 2006; Sanford et al., 2002). However, they would do so because they are sensitive to why the communicator chose that verbal quantifier (as well as why they did not choose another; Sher & McKenzie, 2006). As such, a key philosophical difference lies in the perspective of these two explanations: dual-process theory's framing-effect-as-bias stance argues that people are irrational beings influenced by irrelevant information (Tversky & Kahneman, 1986); pragmatic theory's framing-effect-as-inference stance argues that people are rational beings who use relevant cues to draw reasonable conclusions (Sher & McKenzie, 2008)[4].

Different inferential possibilities may arise depending on the context of the communication (e.g., one could infer the communicator's recommendation; Keren, 2007, the boundary of a range of estimates; Mandel, 2014, or a speaker's reference point; Honda & Yamagishi, 2017; Ingram et al., 2014; McKenzie & Nelson, 2003)[5]. Generally, one may assume that communicators phrase their words to reflect their preferences and recommendations (Teigen & Nikolaisen, 2009). A recipient could thus infer that a quantifier that focuses on positive attributes of a product means the communicator is in favour of a product, since they would use a different quantifier if they were not. Unless one suspects a communicator of being intentionally dishonest (which is not the case in most communicative contexts), one would follow these conventions of conversation to facilitate communication (Grice, 1975). Indeed, there is evidence that speakers

[4]It should be noted that pragmatic theory does not necessarily invalidate dual-process theory. First, it addresses the process of inference drawn from quantifier choice, and not the type of cognitive process used. Second, it does not specify whether the pragmatic inferences drawn are intuitive or analytical. There is a greater literature on whether language inferences, particularly scalar inferences, are automatic or effortful, but they are not the focus of this thesis, and interested readers are referred to Bott & Noveck (2004).

[5]Most pragmatic explanations I discuss in this chapter are not mutually exclusive. People may in fact rely on multiple inferences to influence their decisions. However, the scope of such an investigation would be wider than this thesis could realistically address. In subsequent chapters, I will focus on recommendations from a speaker as a pragmatic inference. How different pragmatic inferences may feed into one another is an unexplored and suggested topic for future study (see Chapter 8).

choose frames to match how attractive they believe options to be (van Buiten & Keren, 2009), to recommend options to listeners (Teigen & Nikolaisen, 2009), and to acknowledge listeners' concerns (Keren, 2007). Ultimately, through the process of shifting attention, directional focus provides cues about what elements of the information are relevant to a decision. These will then have a greater influence on one's judgement and decision-making (Einhorn & Hogarth, 1981).

Challenges for research on pragmatic implicatures with verbal and numerical quantifiers. The challenges in applying pragmatic theory to differences in inferential processes for verbal and numerical quantifiers lie, first, in ensuring equivalence between the quantifier formats (i.e., whether a high chance is indeed a 70% chance for all participants in an experiment), second, in isolating the directional focus of a verbal or numerical quantifier from its contextual framing (i.e., whether '*an uncertain chance of success*' vs. '*a 30% chance of success*' produces the same difference as '*an uncertain chance of failure*' vs. '*a 30% chance of failure*'). The equivalence problem is a similar challenge to that faced by research comparing dual-process explanations for verbal and numerical quantifier differences: previous work has compared the directional focus of verbal quantifiers with those of their average numerical translations (Teigen & Brun, 2000) or membership functions (Juanchich et al., 2013). However, individuals who translate '*a 30% chance*' as a '*moderate chance*' (Welkenhuysen et al., 2001) would have a positive focus compared to individuals who interpret it as '*uncertain*' (Teigen & Brun, 2000), resulting in a more ambiguous directional focus overall. Further, since people tend to overestimate smaller probabilities (Berry, 2006; Budescu et al., 2014), this could explain why numerical probabilities produce more positive directional focus. It is possible that if people's individual interpretations of a negative verbal probability were compared instead, their interpreted numerical probability would indicate an unambiguous negative focus as well. Thus, controlling for individual variation in quantifier interpretation is an important issue for testing the directional focus of verbal and numerical quantifiers.

The second challenge for pragmatic comparisons of verbal and numerical quantifiers is that some verbal quantifiers have a clearly identifiable 'frame' (e.g., the probability expression 'likely' has a clear positive valence, while 'unlikely' has a negative one), while others may possess less inherent valence (e.g., while 'a few' generates a different focus from 'few', it is less clear to a reader which should be considered positive or negative). This is important because a dual-process explanation posits that framing effects rely on the valence of the framing context, while a pragmatic explanation posits that framing effects rely on the directional focus of the quantifier. With probabilities, one can more clearly identify the valence of the quantifier and the valence of the context (e.g., '*an uncertain chance of success*' is a negative quantifier paired with a positive frame), and thus the directional focus of verbal and numerical probabilities can be systematically compared (Teigen & Brun, 2000). Proportional and existential quantifiers, in contrast, are less clear-cut. Would people focus on good or bad applies if '*few of the apples are good*' vs. '*20% of the apples are good*'? Sanford et al. (2002) and Sanford & Moxey (2003) argue that the numerical quantifiers 25% fat vs. 75% fat-free put a different focus on participants' evaluation of the information (75% fat-free is always evaluated for healthiness, but 25% fat may be evaluated from an unhealthy perspective). However, it is possible that the frame itself (fat or fat-free) provides the direction and not the numerical quantifiers. It is therefore necessary to vary frame and quantifier format systematically to isolate the effects of each. Based on the explanation offered by pragmatic theory, we could then expect verbal quantifiers to magnify the framing effect if they indeed possess more inherent and unambiguous directional focus than numerical quantifiers (Teigen & Brun, 2000) and increase pragmatic signals that an action is recommended (Schmeltzer & Hilton, 2014; Teigen & Brun, 2000). In this thesis, I address the two challenges described as follows: Chapter 5 investigates whether proportional verbal quantifiers have a different attention-direction function to numerical ones. Chapters 6-7 provide methodological solutions to the interpretational equivalence problem and systematically compare frame and format of quantifier phrases.

1.4 Practical Applications for Nutrition Communication

Understanding interpretation, attention, and evaluation of verbal and numerical quantifiers through the lens of dual-process and pragmatic theories is useful on an applied level because it informs policy on whether to use a verbal or numerical format to convey quantified information. For example, in the domain of nutrition communication, an important concern for practitioners is which format promotes effective use and understanding of nutrition labels (Malam et al., 2009). In this section, I discuss three practical questions connected to the interpretation, processing style, and framing of nutrition quantifiers. First, can people better interpret a verbal or numerical quantifier? Second, will processing the information in an intuitive or analytical way affect decisions? Third, will people infer implicit details from a communicator's choice of frame? Through this, I show how a systematic, theory-driven empirical investigation can provide crucial input to the applied question of what label format works best, and why.

1.4.1 Implications for interpreting nutrition quantities

As outlined in section 1.2, people tend to interpret verbal quantifiers more vaguely than numerical quantifiers (Budescu & Wallsten, 1985). This could make it more difficult to make precise decisions with verbal quantifiers (Olson & Budescu, 1997). For example, if low % is vaguely interpreted as a range from 10-15%, it would be harder to distinguish two foods that are 10% and 15% fat, which would both be 'low fat'. On the other hand, numerical formats can be falsely precise, and inappropriate, if the quantity involved is uncertain (Budescu et al., 2012). For example, nutrition information is often presented as a percentage of how much one should eat in a day (i.e., daily reference intake or guideline daily amount; Storcksdieck genannt Bonsmann et al., 2010). The reference total is calculated based on average population requirements (e.g., 2000 calories a day), and would vary among individuals (UK Food and Drink Federation, 2009). As such, low % fat might correctly capture the variance in the food's guideline daily amount value for different people.

A more significant problem is that people's interpretations of verbal quantifiers often fall outside the ranges intended by communicators, who tend to underestimate their vagueness (Brun & Teigen, 1988). This misalignment between official and individual translations has been observed in the communication of climate change uncertainty (Budescu et al., 2014), medical side effects risks (Berry, 2006; Webster et al., 2017), and plant pest risks (MacLeod & Pietravalle, 2017), to give several examples. Verbal quantifiers can mean a range of different numerical values to different people (Budescu & Wallsten, 1985). They are also frequently overestimated (Berry, 2006) or interpreted in a regressive fashion (i.e., large verbal probabilities are underestimated and small ones overestimated (Budescu et al., 2012). It is reasonable to extrapolate that such discrepant translations will be present in consumer interpretation of quantifiers used in nutrition labels. However, there has not yet been any empirical work on what numerical values consumers associate with verbal nutrient quantifiers (e.g., 'low') and whether these match official regulations for food labelling.

Three potential implications could arise from misinterpretations of verbal nutrient quantifiers, which are illustrated in Figure 1.1. First, misinterpretations create a comparison problem between products, since label formats are not standardised across manufacturers (Wise, 2013). If one were to compare the two labels in the top panel of Figure 1.1, and a much higher estimate is assigned to verbal quantifiers than intended by guidelines (Berry, 2006), the 'high protein' cereal might seem substantially healthier than the '27% protein' one. In actual fact, its protein content is lower. Second, misinterpreting the numerical amount that a verbal quantifier refers to could hinder consumers' abilities to meet dietary targets. The yoghurt in the middle panel of Figure 1.1 boasts that it is 'high' protein, but this may be interpreted as higher than 30% of one's daily value (the minimum legal requirement to make such a claim; UK Department of Health, 2016). Consumers might then believe they are consuming more protein than the yoghurt actually provides (roughly 20% in this case). Third, different nutrients are often quantified using different formats within the same product, as seen as the descriptions of fibre (verbal) and sugars (numerical) on the cereal in the bottom panel of Figure 1.1. According to labelling guidelines Council of the European Union, 2006, '*high fibre*' should mean at least 30% of one's daily value, but if one overestimates this, one would believe that the cereal contributes substantially more fibre than sugars, when it is not the case. Misinterpretations of verbal nutrient quantities thus impact the quality of people's diets. This issue is addressed in Chapter 2 of this thesis.

Example 1

Protein labels from two different cereals

The verbal label on the left may be mistakenly translated as more than the 27% numerical label on the right.

Example 2

Front and back view of the same yoghurt

The verbal label on the left may be mistaken as more than the actual protein content of the product (right).

Example 3

High fibre (left) and 12% sugar (right) labels from the same cereal

One could believe the cereal contributes much more fibre than sugar.

Figure 1.1. Illustration of problems in translating between verbal and numerical quantifiers in real-world products. Photographs were taken and edited by the author to remove brand labels.

1.4.2 Implications for intuitive and analytical processing of food quantities

Dual-process theory posits that verbal quantifiers could elicit intuitive processing while numerical quantifiers could elicit analytical processing (Windschitl & Wells, 1996). In practical terms, this could mean that a consumer would be more analytical with a '5% fat, 5% fibre' label and more intuitive with a 'low fat, low fibre' one. Based on the characteristics of intuition and analysis, people would evaluate the '5% fat, 5% fibre' food with a slower, more weighted analysis of the quantities of both nutrients (De Neys, 2017b). However, this process would also require more cognitive effort and draw on working memory (Evans & Stanovich, 2013), meaning that people might avoid using numerical nutrient labels to reduce cognitive costs (Peters et al., 2007), or their ability to make effective decisions might be impaired if they are under other cognitive loads (De Neys, 2006), for example, choice overload on a shopping trip (Scammon, 1977; Wansink & Sobal, 2007).

In contrast, people would make quicker decisions about the ‘low fat, low fibre’ food, and rely more on shortcuts to evaluate it (Kahneman, 2011; Tversky & Kahneman, 1974). For example, they might rely on an instinct that low fat foods are healthy (Wansink & Chandon, 2006). Decisions with verbal quantifiers should not overburden cognitive resources, making people more likely to prefer verbal nutrition labels (Malam et al., 2009). However, it could lead to judgement and decision biases, such as being more likely to judge meat that is 25% fat as less healthy than the equivalent 75% lean (Levin & Gaeth, 1988). Understanding whether verbal and numerical quantifiers are intuitively or analytically processed, and how this affects decision-making, thus helps to assess which label format works best for the nutrition labelling context. This issue is addressed in Chapters 3 and 4 of this thesis.

1.4.3 Implications for the framing of food quantities

Pragmatic theory posits that verbal quantifiers create a ‘frame’ by directing the focus of attention to or away from the quantified attribute, and that they do so more than numerical quantifiers (Teigen & Brun, 2000). This could mean that a ‘high fat’ label directs focus to how much fat the food has, but a ‘low fat’ label would direct attention to how much fat the food does not have (Moxey & Sanford, 1986). In contrast, both ‘60% fat’ and ‘5% fat’ should direct focus to how much fat is in the food (Teigen & Brun, 2000). This focus affects how people attend to quantified information, and subsequently evaluate it. Critically, it is the implicit assumptions, or inferences, about the intentions of the communicator and the purpose of the communication that influence people’s judgement (Keren, 2007; Sher & McKenzie, 2006). For example, one might believe the food was labelled ‘low fat’ in order to highlight that one should focus on how little fat the food has. However, if the food were labelled ‘5% fat’, one might believe this was to highlight that one should focus on how much fat it has (Sanford et al., 2002; Teigen & Brun, 2000). This would lead to two different conclusions about how healthy the food is. Therefore, understanding how the focusing properties of verbal and numerical quantifiers frame a food and people’s pragmatic inferences about it can help to assess if consumer's inferences are valid in the nutrition context. This issue is addressed in Chapters 5-7.

1.5 Structure of Thesis

The goal of this research project was to investigate differences in the processing of verbal and numerical quantifiers, with a focus on three key areas: interpretation, attention to, and evaluation of quantified information. In this chapter, I have given an overview of how these issues may be approached, using dual-process and pragmatic theories to provide a guiding theoretical framework. Over the course of this thesis, I will present a systematic investigation over 14 empirical studies that used a range of methods and analyses to address three overarching research questions: (1) What is the nature of the difference between verbal and numerical quantifier processing? (2) How can dual-process and pragmatic theory account for these differences? (3) How do these differences impact judgement and decision-making? I also consider the practical insights offered by answering these questions.

The rest of this thesis is structured as follows:

In Chapter 2, I present Studies 1 and 2, which investigated variability in interpretations of verbal and numerical nutrient quantifiers, whether translations between the formats match standard guideline values, and what predictors account for variations in the magnitude of people's interpretations. The studies were conducted using online survey methods, where participants provided numerical translations of given verbal quantifiers, selected verbal descriptors for given numerical quantifiers, and gave visual scale estimates for both types of quantifiers. These methods were adapted from previous research in interpretations of verbal probabilities (Du & Stevens, 2011; Theil, 2002).

In Chapter 3, I present Studies 3 and 4, which investigated the dual-process mechanisms for processing quantifiers using a nutrient quantity decision task. This task was designed to test the basic tenets of dual-process theory[6]:

[6]Many dual-process theories exist, with variations on what combinations of attributes dis-

intuitive processes proceed quicker, with less effort, and result in sub-optimal performance compared to analytical processes. Participants had to integrate a given nutrient quantity (presented pictorially to avoid priming verbal or numerical processing) with a new quantity (either verbal or numerical) and decide if it would exceed their recommended daily total consumption of the nutrient. These experiments were delivered in the lab using average translations of verbal quantifiers obtained in Chapter 2 (Study 3) and online using participants' own translations for numerical quantifiers (Study 4).

In Chapter 4, I present Studies 5 and 6, which built on the studies in Chapter 3 to provide a critical test of another postulate of dual-process theory: analytical processes are affected by concurrent cognitive loads, while intuitive processes can continue to operate normally under load. We used a dot memory task to interfere with analytical processes (Białek & De Neys, 2017) and combined this with the decision task from Chapter 3. These experiments were also delivered in the lab using average translations of verbal quantifiers obtained in Chapter 2 (Study 5) and online using participants' own translations for numerical quantifiers (Study 6).

In Chapter 5, I present Study 7, which used eye-tracking methodology to trace attention processes when participants observed verbal or numerical nutrition labels. In this study, we compared two hypotheses derived from the dual-process and pragmatic theories respectively. Based on dual-process theory, we would expect participants to focus more attention on numerical labels if they elicit more analysis than verbal ones. Based on pragmatic theory, we would expect participants to focus more attention to different components of the label (e.g., the nutrient or the quantifier) between verbal and numerical labels. We further investigated using mediation analyses whether differences in attention would explain variations in participants' judgements of the nutrition label.

In Chapter 6, I present Studies 8 to 10, which investigated the directional focus of verbal and numerical quantifiers using the attribute framing effect (Levin

tinguish intuitive and analytical processes. We used the most basic form of the theory to derive our hypotheses. There are other ways to approach the issue, which I address in later chapters.

& Gaeth, 1988). This paradigm presents participants with a vignette about a beef that is either 75% lean (positive attribute frame) or 25% fat (negative attribute frame). Participants judge the beef on various dimensions (e.g., preference, healthiness). Study 8 investigated whether verbal quantifiers would produce a magnified framing effect for judgements of healthiness for the beef compared to numerical quantifiers, over four combinations of quantities. Studies 9 and 10 tested the robustness of the format effect and whether participants would be directed by the frame and format to focus on either the absence or presence of the attribute (fat or lean meat). These studies controlled for variability of quantifier translations by using a translation procedure whereby participants selected verbal quantifiers to match the numerical quantifiers in the constructed equivalence frames.

In Chapter 7, I present Studies 11 to 14, which built on the studies in Chapter 6 to compare the dual-process and pragmatic explanations for different quantifier formats, using a modified framing paradigm. We used the concept of food energy to create a minimal framing paradigm that controlled for variations in interpretations of quantifiers of different magnitudes: '40% energy' is equivalent to '40% calories'. This kept the quantifier constant despite a change in the attribute's valence. Studies 11 and 12 measured participants' affective associations with the frames ('energy' and 'calories') and conducted a moderated mediation analysis to ascertain the role of affect in explaining the attribute framing effect as modified by quantifier format. Study 11 tested this paradigm using averagely translated numerical values, while Study 12 used participants' self-produced numerical translations to account for interpretational variability. Study 13 used the same procedure, but tested for participants' inferences about what a speaker's recommendation was based on the frame. We used this instead as the mediator in the statistical analysis for this study. Study 14 compared both affective associations and inferred recommendations in a single integrated model to explain the attribute framing effect across different quantifier formats and magnitudes.

Finally, Chapter 8 summarises the results of all my studies and provides a discussion of the theoretical and practical implications of the findings.

1.6 Open Science Contribution

1.6.1 The replication crisis

Research in psychological science currently faces a challenge in producing replicable findings (Open Science Collaboration, 2014; Stevens, 2017; Yong, 2012). For example, a recent investigation found that less than half of 100 selected psychology studies could be replicated, with most of the original effect sizes substantially smaller than originally reported (Open Science Collaboration, 2015). This large-scale project addressed a growing concern in the field that the literature is populated by an abundance of false positives (Ioannidis, 2005; Maxwell et al., 2015; Stevens, 2017).

The structures that perpetuate the replication crisis are manifold. Conducting a replication can be a challenge as original materials and data are often unavailable due to information loss (Open Science Collaboration, 2014) or unwillingness of the original researchers to share published data (Wicherts et al., 2011). Further, publication bias, in which significant and consistent findings are favoured over non-significant and inconsistent ones, creates a barrier for reporting replication studies that fail to reproduce the original result (Makel et al., 2012; Open Science Collaboration, 2014; Renkewitz et al., 2011). This incentivises researchers to focus on 'newsworthy' and 'eye-catching' findings (Laws, 2016; Yong, 2012), and engage in questionable research practices that increase the likelihood of finding a significant (and thus, 'publishable') result (John et al., 2012). For example, researchers could base their sample size on whether they achieved a significant result with the existing sample, or even generate and report a hypothesis after the results are known, thereby ensuring that results are in line with one's hypotheses (John et al., 2012). Scientific research practices thus need to take clear measures to tackle these threats to the reproducibility and robustness of research findings (Munafò et al., 2017).

1.6.2 Improving the reproducibility of research through open science

Open science, which encourages transparency and openness about one's research process, provides a framework to improve research reproducibility (Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J., ...Yarkoni, T., 2015). Guidelines for open science practice provide several recommendations to overcome the threats to reproducibility outlined above (Open Science Collaboration, 2014). An overarching recommendation is to conduct pre-registered research. This involves specifying study protocols, research objectives and hypotheses, and data collection and analysis plans prior to beginning a study (Munafò et al., 2017; Open Science Collaboration, 2014). Pre-registration introduces a level of accountability for researchers to adhere to and prevents questionable decisions such as choosing analyses depending on the nature of the data collected (Munafò et al., 2017). Further, it can also encompass other elements that help to improve reproducibility:

Data sharing. Pre-registering a study often involves a commitment to make public the study results. Making available research materials and data allows future replication efforts to be conducted; more importantly, it helps to combat publication bias by providing platforms for sharing scientific research regardless of the outcome (Munafò et al., 2017).

Statistical power. When research is conducting using low statistical power, it runs the risk of being unable to detect a small effect, but also that a significant result may not reflect a true effect (Button et al., 2013). Standard pre-registration protocols should specify a stopping rule for data collection. This encourages researchers to carefully consider the target sample size and conduct suitable *a priori* power analyses.

Confirmatory hypothesis testing. A pre-registered study should specify the *a priori* analysis plan, thus distinguishing between confirmatory and exploratory research (Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J., ...Yarkoni, T., 2015). Performing confirmatory analyses first in accordance with the pre-registered plan safeguards against the possibility of conducting multiple analyses to turn up a significant result, or hypothesising after the results are known (Munafò et al., 2017). This improves reproducibility by clearly identifying when a specified effect is being tested and when an effect may contribute to a subsequent hypothesis.

1.6.3 This thesis's contribution to open science practices

Motivated by the need to contribute to better scientific practice in psychological research, the work in this thesis was informed by the principles of open science. Adhering to the recommended open science practices discussed above presented several challenges; I address these, along with reflections on how to facilitate open science for doctoral research, in Chapter 8.

All studies from Chapter 3 onwards were pre-registered with the specific hypotheses, analysis plan, and stopping rules for data collection for each study determined *a priori*. Where possible, I determined sample size by power analysis based on known effect sizes, or those obtained in the earlier studies. In cases where *a priori* stopping rules were determined based on time constraints, I conducted sensitivity analyses to report the estimated power of the collected sample to detect the given effect.

To facilitate open sharing of research resources, I have archived all the pre-registrations, material, and data reported in the studies on the Open Science Framework (OSF). Links to the specific data repository for each study are included in the respective chapters.

Chapter 2: People Overestimate Verbal Quantities of Nutrients on Nutrition Labels

2.1 Abstract

Nutrition labels provide information about nutrient quantities in food, thus offering consumers a tool to make healthy eating choices. These labels are often presented with verbal quantifiers (e.g., 'low'). However, little is known about how consumers actually interpret this information. We investigated how participants understand verbal quantifiers, whether these translations fit standard guidelines, and whether nutrient valence and individual differences predicted interpretational variability. In Experiment 1 ($N = 82$), participants gave numerical percentages for five verbal quantifiers, selected a verbal expression that best described eight numerical quantifiers, and estimated the amount conveyed by quantifiers of both formats on a visual analogue scale. In Experiment 2 ($N = 801$), participants translated five verbal quantifiers into numerical percentages. Participants interpreted quantifiers with great variability and substantially overestimated the numerical value of verbal quantifiers as compared to standard guidelines. The magnitude of estimations persisted across participants with different individual characteristics. It may be beneficial to refine guideline ranges for nutrient values to better match people's intuitive interpretations of verbal quantifiers.

2.2 Introduction

From an individual and a public health perspective, the ability to judge the healthiness of food is important for combating the increase in diet-related disease worldwide (Crockett et al., 2018). To improve dietary behaviour and curb rising obesity rates (Craig & Mindell, 2014), public health organisations worldwide have spearheaded legislation for nutrition labelling to increase people's awareness of the nutrient content of the food they are eating (e.g., US Food and Drug Administration, 2004; Wise, 2013). Nutrition labelling is intended to empower people to make informed and healthier choices in their food purchases and consumption (Crockett et al., 2018; Hawley et al., 2013). However, nutrition labels can sometimes be cryptic and misunderstood. How healthy is a '15% fat' product, for example? Is it healthier than a 'low fat' product? In the present study, we examined whether people interpret nutrient quantities in nutrition labels appropriately.

Verbal descriptors were introduced to simplify and categorise numerical quantities on nutrition labels. Specifically, the terms 'low', 'medium', and 'high' were introduced as verbal banding to provide consumers with a quicker and more intuitive understanding of nutrient amounts (Malam et al., 2009). Official guidelines specify what numerical ranges verbal labels should correspond with (see Table 2.1 for an example from the UK; UK Department of Health, 2016), as well as when such terms (e.g., 'high in minerals') can be used as nutrition claims by manufacturers (UK Department of Health, 2011).

Table 2.1. Typical Guideline Daily Amount (GDA) Ranges Associated with Verbal Descriptors on Nutrition Labels

Nutrient	Low % GDA	Medium %GDA	High %GDA
Fat	$\leq$ 4.3%	4.3-25%	>25%
Saturates	7.5%	7.5-25%	>25%
Sugars	5.6%	5.6-25%	>25%
Salt	5%	5-25%	>25%

Note. Advice to consumers on how to interpret verbal quantities is determined in terms of amount per 100g, although they are often presented with GDA labels that present the percentage contribution of a nutrient to total recommended daily intake. To simplify presentation, we have calculated the equivalent percentages using the guideline amount per 100g. The necessity of doing so further demonstrates the complexity involved in understanding guidance information on interpreting GDA information.

Three issues bearing theoretical and practical implications are relevant to the interpretation of verbal quantifiers. First, what do verbal quantities like ‘low’, ‘medium’, and ‘high’ mean to consumers, and how much do they vary across people and contexts? Second, do these interpretations match the intended meaning postulated in standard guidelines? Third, what factors contribute to variations in interpretations of verbal quantifiers? Answering these questions would shed light on whether consumers are actually making informed food choices. To address these issues, we drew on theoretical and empirical advancement from previous work on the interpretation of verbal expressions of risk and frequencies.

2.2.1 How low is ‘low’? Average interpretations of verbal quantities and their variability

There is no direct evidence about how people interpret verbally communicated food quantities, however discrepancies in people’s interpretations of verbal quantifiers have been demonstrated in domains such as medical side effects (Berry et al., 2003) and climate change outcome likelihoods (Budescu et al., 2014). People can derive different interpretations of a same verbal probability depending on the context and their knowledge of the subject (Harris et al., 2009; Knapp et al., 2015; Sirota & Juanchich, 2015). This affects their perceptions of a quantity, and their subsequent decisions (Berry et al., 2003; Juanchich et al., 2012). It is therefore important to ascertain what ranges people actually ascribe to verbal quantities such as ‘low’, ‘medium’, and ‘high’, rather than assuming a consistent and accurate interpretation of these terms.

2.2.2 Lay interpretations of verbal quantities may not match standard guidelines

Studies on how people translate from words to numbers in verbal expressions of uncertainty have found that people tend to interpret verbal probabilities in a regressive fashion, with their numerical translations clustering around the centre of a numerical scale (Budescu et al., 2014). Furthermore, when official numerical translations are smaller than people’s intuitive understanding of the verbal term, people are likely to overestimate verbal quantities. This is observed with verbal risk frequencies for medical side effects: the EU assigns a frequency of 1-10% for ‘common’ but respondents believed it to mean 44% on average (Berry, 2006). Verbal banding for nutrition labelling likewise focuses on quantities below 50%, with cut-offs for 'high' amounts around 25-30%. People might thus overestimate nutrient quantities depicted by verbal labels compared to official standards. This could contribute to suboptimal decisions about consumption.

One might argue that misinterpretation of verbal nutrient quantifiers is a trivial problem since verbal labels should enable people to easily choose the healthier of two products without performing any translations (van Herpen & van Trijp, 2011). However, unlike in studies that investigate the effect of label format on food choice, people can encounter two products that each have a different label format because formats for interpretative labels are left to the prerogative of the manufacturer (Wise, 2013). If an individual must decide between a cereal with 'medium' fat and another with '30%' fat, the latter cereal, despite having more fat, could be mistaken as healthier if 'medium' were interpreted as more than 30%. As such, errors in translating from a verbal quantifier to a numerical one can result in people making choices that are less healthy than they expect.

To further illustrate the problem a misinterpretation of verbal quantifiers could pose, imagine that a parent wishes to decide if a cereal with 'high minerals' (which may only have 30% of one's daily requirement of minerals) provides substantial nutrition for their child. If a higher estimate were assigned to the verbal quantifier, one would overestimate the amount of minerals the cereal contributes. Even if verbal labels can facilitate a choice between the healthier of two cereals, they could still result in a misjudgement of the actual nutritional value of the food.

2.2.3 Can interpretational variation be predicted?

People draw on prior knowledge, beliefs, and attitudes to interpret information. A motivated reasoning account of quantifier interpretation suggests that subjective biases such as the desirability of a particular outcome can influence the interpretation of quantifiers (Kunda, 1990). Factors such as the valence of the nutrient described, familiarity-based recognition of labels, or self-serving biases could account for variability in the perception of the same food quantity. A nutrient's valence can cue the severity of outcomes associated with its consumption (Oakes, 2005b). Severe outcomes may be perceived as having greater likelihood than neutral outcomes with the same base rate (Harris et al., 2009). We could therefore expect people to provide higher estimates for 'negative' nutrients that indicate the unhealthiness of food than for 'positive' nutrients that indicate healthiness. Alternatively, people might wish to believe their food to be higher in desirable quantities than undesirable ones (Kunda, 1990), and provide higher estimates for positive than negative nutrients.

Individual differences among people can also predispose them to over- or underestimating verbal quantifiers (Moxey & Sanford, 1992). For instance, two people may interpret a verbal frequency of 'high physical activity' as meaning different amounts of time depending on whether they like to exercise or not (Moxey & Sanford, 1992). Prior experience with nutrition labels and motivations for healthy eating could therefore affect how a particular individual interprets them. A person who frequently consults nutrition labels may be more familiar with industry standards in translating verbal quantities, resulting in more accurate interpretations (Gigerenzer et al., 2005). Additionally, the desire to draw a certain conclusion (Piercey, 2009) or support a particular worldview (Budescu et al., 2012) may influence the translation of a verbal quantifier. An individual who values healthy eating might be more motivated to justify their choices as healthy (Rayner et al., 2001), and therefore translate quantifiers in a way that fits this motivation (e.g., reassuring themselves that a 'high %' of a 'negative' nutrient is a lower numerical value that still fits within their daily limit, and vice versa). In this case, motivations to be healthy would prompt people to underestimate quantities of unhealthy nutrients and overestimate quantities of healthy ones.

2.2.4 The present research

The two experiments presented here had three aims. First, we aimed to explore what were people's actual interpretations of verbal nutrient quantifiers. Based on previous findings for probability quantifiers and in different domains, we extended to the food domain the hypothesis that participants' numerical estimates of verbal quantifiers would vary widely between individuals. Second, we aimed to find out whether and to what extent these interpretations fell outside the range of recommended guidelines. We expected to replicate findings from the climate change and medical domains that the interpretations would be misaligned with official interpretations. Finally, we explored factors that might moderate variability in people's perceptions of verbal and numerical nutrient quantifiers, and if people's perceptions would persist across product or individual characteristics. Drawing from the literature on motivated reasoning (Kunda, 1990) and severity effects (Harris et al., 2009), we derived two possible hypotheses for how product characteristics (positive or negative nutrients) would affect people's perceptions: either participants would provide either higher estimates for positive than negative nutrients, or vice versa. This provided a novel test of which of the two accounts might apply to interpretations of nutrient quantifiers. Based on research on the effect of individual differences in motivation and knowledge in quantifier interpretation (e.g.,Budescu et al., 2012; Gigerenzer et al., 2005), we expected to replicate these effects in our quantifier domain: participants for whom healthy eating was important would be motivated to provide higher estimates for positive nutrients and lower ones for negative nutrients.

To address these aims, we used translation tasks that were widely used in the literature on verbal probabilities (Collins & Hahn, 2018) to test participants' interpretations of verbal nutrition quantifiers. While this method allowed us to solicit interpretations of quantifiers along a mixed scale (0-100%), it required that we standardise values of nutrients. Therefore, we tested participants' interpretations in terms of the percentage of their reference intake provided by a food (or 'Guideline Daily Amount'; hereafter 'GDA'; UK Food and Drink Federation, 2009). GDA labels are the most widespread among front-of-package food labels across the EU (Storcksdieck genannt Bonsmann et al., 2010) and are most viewed when consulting nutrition labels (Grunert et al., 2010a). Although official verbal quantifier ranges are determined in terms of amount per 100g (see footnotes to Table 2.1), verbal quantifiers are often presented with GDA percentages, and we believe that consumers could interpret them accordingly.

We report exploratory data for our three study objectives in Experiment 1, and tests of the hypotheses on a larger sample in Experiment 2. All the data, materials, and supplementary analyses reported in text are available on the Open

Science Framework (OSF).

2.3 Experiment 1

2.3.1 Method

Participants. Participants (N = 82; 83% female; age range 18-66, M = 21.41, SD = 7.90) were recruited from a volunteer list of a UK university. The sample size was determined by an a priori time-based stopping rule. A post-hoc sensitivity analysis revealed an ability to detect an effect size of f = .31 with 80% power (α= .05). Undergraduate students were given module credits for participation. Participation for non-students was voluntary. The sample was 53% White, 16% Asian, and 20% African (11% other races); 45% had a university degree. Participants had a healthy estimated average Body Mass Index (BMI; M = 22.04, SD = 5.31), reported average attitudes towards healthy eating (M = 4.89, SD = 0.98 on a 7-point scale with higher values reflecting more positive attitudes) and 48% agreed that they frequently looked at nutrition labels.

Design. In a within-subject experiment, participants interpreted the meaning of 13 GDA labels that were presented with a range of nutrients (see Figure 2.1). Participants indicated how important these nutrients were in determining the healthiness of food. This provided an indication of each nutrient's relative valence. We also measured participants' attitudes towards healthy eating, frequency of nutrition label use, and estimated their BMI from their given height and weight ranges.

Materials and procedure. Participants completed the experiment online. After providing informed consent, they completed a healthiness ranking task, a set of interpretation tasks, and a set of individual difference measures, in a fixed order as presented here.

Healthiness ranking task. Participants ranked a list of eight randomly ordered nutrients according to their importance in determining (1) the healthiness and (2) the unhealthiness of food. We used the selected rankings for this task to assign the verbal and numerical quantities participants would see in conjunction with a nutrient. This ensured that participants saw all nutrients and all values, but were not overwhelmed by having to rate all possible combinations. Based on the mean ranks of nutrients (reported in A, Table A.1), we also categorised the four most important nutrients for determining healthiness as ‘positive’ (minerals, protein, calories, and fibre) and the four most important for determining unhealthiness as ‘negative’ (sugar, fat, sodium, and saturates).

Interpretation tasks. Participants interpreted nutrient labels showing a single percentage of the GDA for eight nutrients. Three interpretation tasks were presented in a random order for each participant (see Figure 2.1). The quantifiers in each appeared with a different nutrient, and the assignment of quantifier to nutrient varied across participants. Within each task, the labels were presented separately (one label per page) and in a randomised order for each participant. After each task, participants rated how easy they found the task on a five-point Likert scale (1: extremely difficult, 5: extremely easy).

Numerical translations of verbal labels. Participants translated five verbal labels (see Table 2.2) by answering the question, ‘*What percentage of a guideline daily amount (GDA) of [nutrient] do you think the food label describes? (Please give a number.)*’

Back-translations of numerical labels. Participants matched eight numerical labels to verbal quantities. To provide their answers, they selected from a multiple-choice list of five verbal quantifiers (*very low, low, medium, high, very high*). We focused our analysis on the five numerical quantifiers that best matched interpretations of the five verbal quantifiers (see Table 2.2). Participants’ interpretations of the other numerical quantifiers (10%, 75%, and 90%) are included in A (Table A.3).

Quantity perceptions. Participants estimated the GDA proportion described by the specified quantity for the 13 verbal and numerical labels. They gave their estimates on a visual analogue scale with three anchor points (*none, half*, and *all*) corresponding to 0-100 with invisible increments of one. This visual analogue proportion measure allowed a standard comparison between participants’ perceptions of verbal and numerical quantifiers. We could therefore determine if there was a convergence between their perceptions of verbal and numerical quantifiers.

Individual difference measures.

Attitudes towards healthy eating. Participants completed the healthy eating motivation section of the Food Choice Questionnaire (Steptoe et al., 1995). This included seven questions answered on a 7-point Likert scale (1: *strongly disagree*, 7: *strongly agree*); for example, '*I always follow a healthy and balanced diet.*' The scale showed good reliability (Cronbach's α= .78). Participants' attitude towards healthy eating was computed as the average of scores, with higher scores reflecting more positive attitudes.

Frequency of nutrition label use. Participants answered an additional statement included in the healthy eating attitude scale: '*I often use nutritional labels to determine the healthiness of food*'.

Estimated BMI. Participants provided estimates of their weight and height by selecting from a drop-down list of six weight and height ranges. Estimated BMI was calculated for each participant by taking the middle value of the weight range (in kilograms) divided by the square of the middle value of the height range (in metres).

Socio-demographic characteristics. Participants reported their age, gender, ethnicity, and highest level of education completed.

2.3.2 Results

Participants' interpretations of verbal quantifiers. To determine people's actual interpretations of verbal quantifiers, we first analysed the numerical percentages that participants assigned to each verbal label and the back-translations of the verbal descriptors assigned to numerical labels. Participants' numerical estimates of verbal labels varied to a large extent; for example, from 5-100% for 'high' (see Table 2.2. A repeated measures ANOVA showed that there was a significant difference between the numerical values provided for the different verbal quantifiers, $F(4, 292) = 322.89$, $p < .001$, $\eta^2_{P} = .82$. The back-translation of numerical quantifiers into verbal ones generally matched the mean numerical

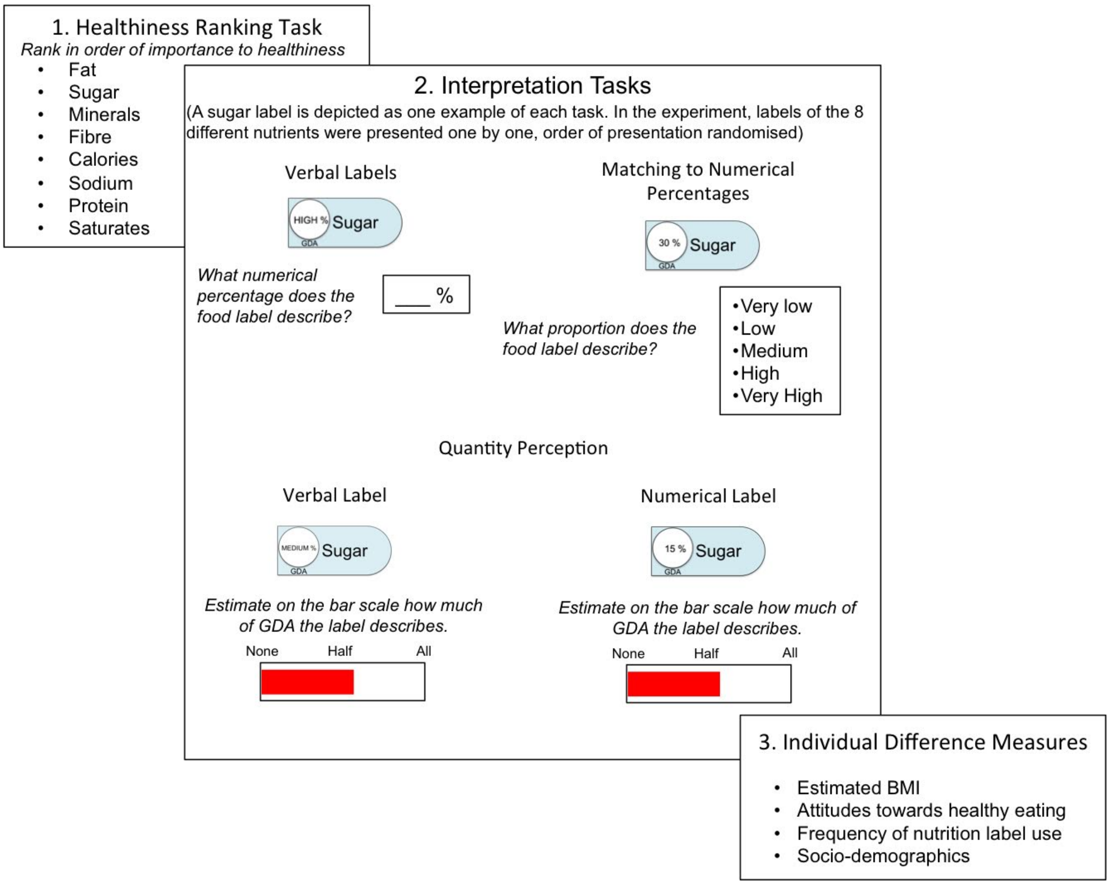

Figure 2.1. Structure of experimental tasks in Experiment 1.

Note. A sugar label is depicted as an example. In Experiment 1, each participant saw a different nutrient for each quantifier. In Experiment 2, only the verbal label interpretation task was used, and participants saw either fat or minerals only.

estimates given for the translation of verbal into numerical quantifiers. The numerical values were consistently ranked in ascending order when assigned back to verbal quantifiers, Kendall's w = .62, χ^2 (7, N = 75) = 325.25, p < .001. The perceptions of both verbal and numerical quantifiers on a visual analogue scale showed great variability (see Figure 2.2).

Do participants' interpretations match standard guidelines? Since our sample was British, we compared participants' interpretations of verbal quantifiers to available guideline ranges in the UK for negative and positive nutrients. As anticipated, participants calibrated their estimates to the full range of a per-

Table 2.2. Interpretations of verbal and numerical quantifiers in Experiments 1 and 2

Experiment 1(N = 82)					
	Numerical interpretation of verbal labels			**Verbal interpretation of numerical labels**	
Label	***M***	***SD***	**Range**	**Label**	**Median**
Very Low %	9.36%	12.35	1.00-70.00	10 %	Low
Low %	16.92%	16.10	1.00-90.00	15 %	Low
Medium %	43.13%	12.08	15.00-70.00	50%	Medium
High %	68.16%	19.62	5.00-100.00	75%	High
Very High %	78.24%	20.50	10.00-100.00	90%	Very High

Experiment 2 (N = 801)						
Label	**Fat**		**Minerals**		**Average**	
	M	***SD***	***M***	***SD***	***M***	***SD***
Very Low %	7.54%	11.56	12.69%	17.32	10.11%	14.93
Low %	11.36%	12.47	16.46%	15.74	13.90%	14.42
Medium %	23.29%	16.63	37.48%	19.11	30.36%	19.26
High %	40.83%	27.37	57.81%	28.15	49.29%	29.02
Very High %	48.93%	30.51	66.15%	29.70	57.51%	31.30

centage scale, resulting in numerical estimates that were much higher than the typical ranges in recommended guidelines, with 56-88% of interpretations exceeding recommended low and medium ranges (see the top panel of Figure 2.3). While most estimates for high values tended to fall in the correct range by default, it should be noted that for positive nutrients, the 30% cut-off is a minimum value, below which the declaration of 'high' cannot be legally used. Therefore 88% of participants were overestimating these quantities of positive nutrients.

Predictors of variability in interpretations. Because each participant viewed different nutrients with different quantifiers, we used separate ANCOVAs to analyse the interpretations of each verbal quantifiers. In the ANCOVA, nutrient valence was used as a factor, and BMI, eating attitudes, frequency of

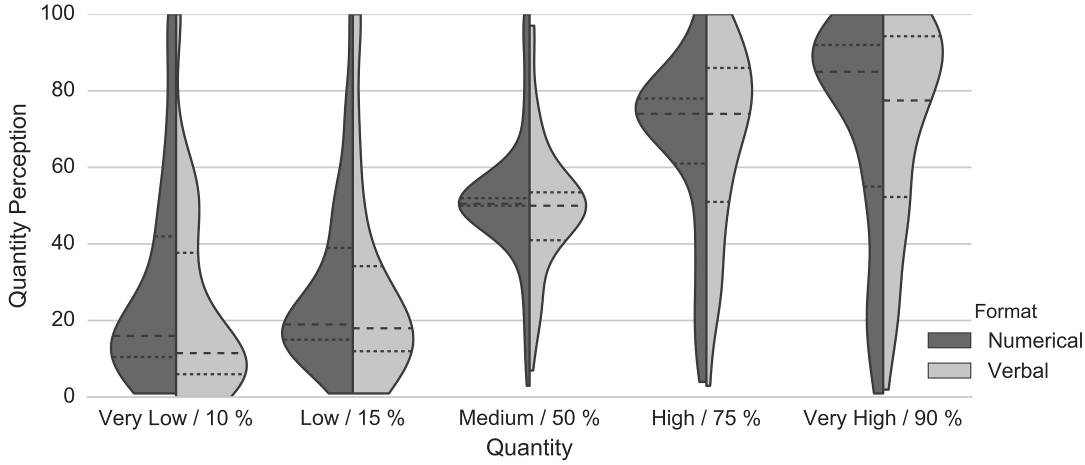

Figure 2.2. Distribution of estimates on the visual analogue scale for verbal and numerical GDA labels. Dotted lines indicate the medians and inter-quartile ranges.

Note. The violin plots show smoothed distributions of the values in the data for each condition, with shading indicating the probability density of each value. Peaks in the plot indicate values would be represented with the highest probability in the population.

label use, age, gender, level of education, and ethnicity were used as covariates. We also examined if attitudes, BMI, and label use would interact with nutrient valence. Full results of the ANCOVA are provided in A (Table A.5). None of these factors or covariates affected the interpretation of labels, all ps $> .05$.

2.3.3 Discussion

Experiment 1 showed that participants' translations of verbal quantifiers varied greatly and did not match standard guidelines. The extent of overestimation was not attributable to nutrient valence or individual differences. However, a number of features of the design may have limited the potential to detect these effects, so we conducted a second replication experiment on a larger sample to investigate this possibility.

First, Experiment 2 recruited a general population sample, to address the limitation of Experiment 1 in having a relatively homogeneous and well-educated

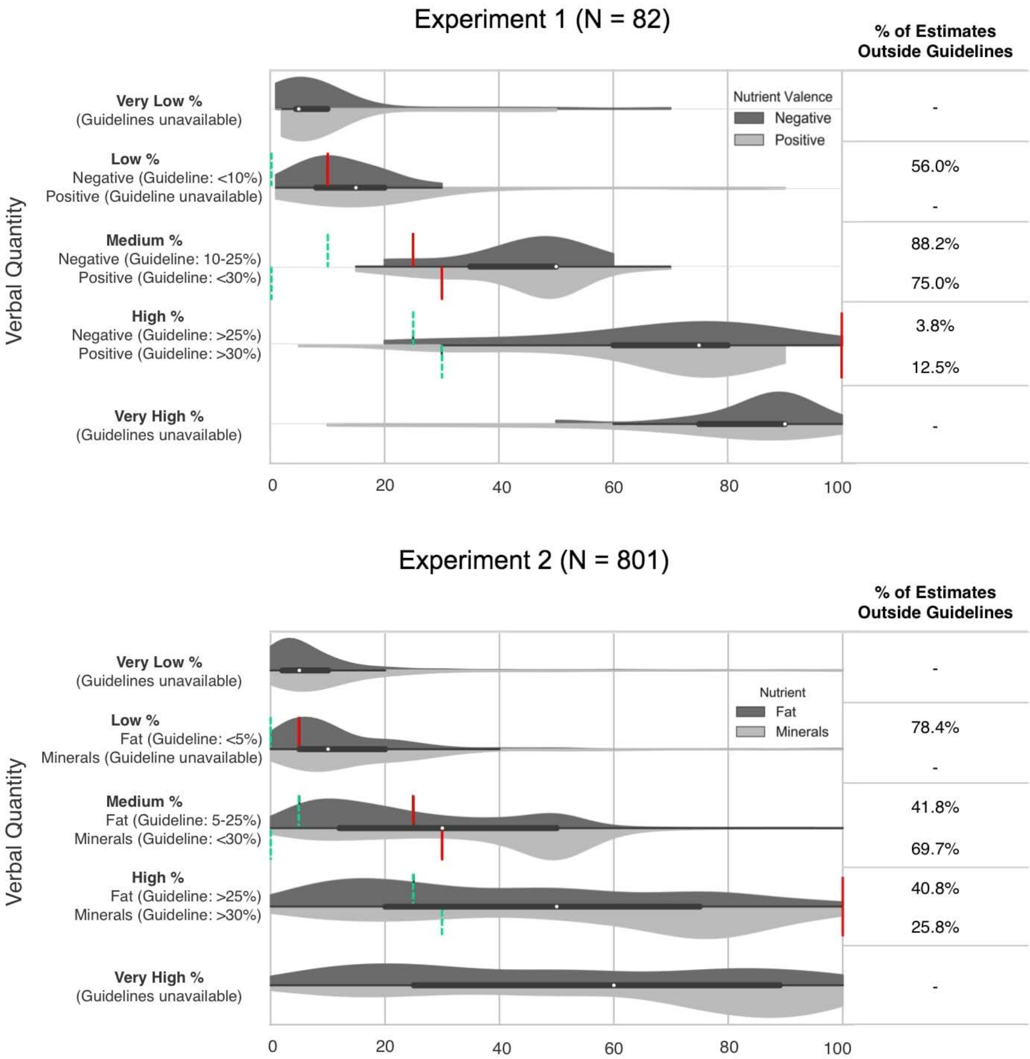

Figure 2.3. Distribution of verbal quantifier interpretations in Experiments 1 (N = 82) and 2 (N = 801), and their fit with recommended guidelines (solid red lines indicate upper limits and dotted green lines indicate lower limits). Guideline values were estimated from the UK Department of Health (2011, 2016). For positive nutrients, only guidelines marking the boundary between medium and high % were available.

Note. The violin plots show smoothed distributions of the values in the data for each condition, with shading indicating the probability density of each value. Peaks in the plot indicate values would be represented with the highest probability in the population.

sample that was likely to have reasonable understanding of GDA labels (Grunert et al., 2010b) and familiarity with general nutrition guidelines (Blitstein & Evans, 2006; Parmenter et al., 2000), resulting in fewer differences between individuals. Second, we amended the simplified BMI measure in Experiment 1 (which took the average of participants' selected weight range) that might have limited our analysis of this individual difference variable. Finally, we used a between-subjects design that assigned participants to a nutrient that had been clearly categorised as healthy (minerals) or unhealthy (fat). This addressed the possibility that the within-subject design of Experiment 1 might have resulted in participants anchoring their subsequent interpretations to the first nutrient they saw.

2.4 Experiment 2

2.4.1 Method

Participants. We sourced a nationwide sample of participants from a survey panel, using quota sampling to determine the demographics of the sample. After excluding participants who did not complete the survey and those who responded carelessly (e.g., giving equal responses to all questions), there were 801 participants (52% female; age range 18-74). This sample was determined by an *a priori* stopping rule and had a post-hoc sensitivity to detect an effect size of f = .10 (α= .05, 1-β= .80). Full socio-demographic characteristics for our participants are in A (Table A.7) . Compared to Experiment 1, there was a larger range of ages and education levels, but reported eating attitudes and nutrition label use were similar ($M_{\text{attitudes}} = 4.65$, $SD = 1.09$; 49% reported frequent use of nutrition labels). Mean estimated BMI was higher than Experiment 1 ($M = 27.36$, $SD = 6.67$).

Design, materials, and procedure. Participants first gave informed consent and provided sociodemographic information. We simplified the verbal label interpretation task from Experiment 1 such that participants translated the five verbal labels for only one nutrient, which was either fat or minerals (randomly allocated). They completed the task following an unrelated questionnaire investigating feelings about seeing clusters of holes. At the end of the survey, participants provided their weight and height in kilograms and centimetres. They also completed a reduced, four-question version of the attitudes towards healthy eating measure (Steptoe et al., 1995; Cronbach's α= .72) and indicated their frequency of nutrition label use.

2.4.2 Results

Consistent with Experiment 1, participants' numerical interpretations of verbal labels displayed large variances (see Table 2.2 and the lower panel of Figure 2.3). There was a significant difference between the numerical values provided for the different quantifiers, $F(4, 3196) = 1283.66$, $p < .001$, $\eta_P = .62$. We provide the pairwise comparisons between quantifiers and nutrients in Appendix A (Tables A.8 and A.9).

Similar to Experiment 1, we found extensive overestimation of the meaning of verbal quantifiers, with 35-88% of low and medium estimates falling above guidelines (see the lower panel of Figure 2.3). Although estimates for high % were technically within the guideline range, nearly half our participants interpreted high % fat GDA as more than the recommended 25%, and three-quarters interpreted a high % mineral GDA in excess of the legal 30% minimum requirement. If participants' interpretations were what they believed to be the cut-off points, they might be substantially overestimating the legal minimums. Numerical estimates for minerals were significantly higher than those for fat, $F(1, 799) = 104.54$, $p <$ .001, $\eta^2{}_P = .01$.

To examine the effect of individual differences on interpretations, we ran an ANCOVA (full results in A, Table A.6). BMI eating attitudes, frequency of label use, gender, age, education level, ethnicity, occupation, and whether the participant as a native English speaker did not predict participants' interpretations of verbal labels, all ps > .18. We also analysed interactions between nutrient valence and the following covariates: BMI, eating attitudes, frequency of label use, but obtained no significant interactions, all ps > .14.

2.4.3 Discussion

Experiment 2 replicated the pattern of results in Experiment 1 over a large sample that included participants from differing demographic backgrounds. There was great variability in the interpretations of verbal quantifiers, many of which fell outside their intended ranges. In addition, we found a significant effect of nutrient valence. These effects were not moderated by individual differences between participants.

2.5 General Discussion

Findings from two experiments showed that the interpretation of verbal labels varied greatly between individuals and that on average, these interpretations did not match their intended meaning. This is the first study to extend this phenomenon, previously found for probabilities (e.g., Budescu et al., 2012) and frequencies (e.g., Berry et al., 2003), to the context of nutrition labelling. We showed that participants consistently overestimated the numerical value that a verbal quantity of nutrients should refer to. Experiment 2 also showed that participants gave higher quantity estimates for a positive nutrient (i.e., minerals) than a negative one (i.e., fat). However, the variability of interpretation of verbal quantifiers persisted across individuals with different characteristics (e.g., BMI, attitudes to healthy eating, and familiarity with labels).

2.5.1 Interpretations of verbal quantifiers vary across individuals

We showed that people interpreted verbal nutrient quantifiers with large variations. For example, a low% indicated 5% of the GDA for one participant, but 40% for another. These findings complement previous research that demonstrates great variability in the mapping of verbal probabilities onto numbers (Berry et al., 2003; Budescu et al., 2014; Budescu & Wallsten, 1985). We extended previous work on the interpretational vagueness of verbal probabilities to verbal quantifiers in the context of nutrition labelling. This replicates findings from other fields, such as for communicating side effect frequencies (Berry et al., 2002), climate change uncertainty (Budescu et al., 2012), and environmental risks (MacLeod &

Pietravalle, 2017).

Given the large variation among individuals in interpreting verbal quantifiers, the advisability of using verbal labels seems questionable. However, verbal expressions remain widely used because they are more natural in communication (Zimmer, 1983) and do not require precise estimation of quantities. For instance, GDA information is meant as a recommended guideline based on population averages for daily required nutrient intakes (UK Food and Drink Federation, 2009); it would be inaccurate to assume a 5% GDA label indicates exactly 5% of one's GDA. Here, a 'low %', thanks to its vagueness, may actually convey the information more accurately, provided its range is interpreted as intended.

2.5.2 Interpretations of verbal nutrient quantities are largely above standard guidelines

Critically, interpretations of verbal quantifiers were much higher than the values in recommended guidelines, indicating that people overestimated the quantity of nutrients a food provided. Similar to findings for medical side effects (Berry et al., 2002, 2003; Webster et al., 2017), but in contrast to those regarding climate change probability perceptions (Budescu et al., 2012, 2014), participants tended to overestimate the numerical value of all the verbal quantifiers. This could be because people naturally anchor verbal quantifiers to the full range of a percentage scale instead of the ranges provided by food labelling guidelines (which are skewed low, from 5-30%). A similar tendency is observed for overestimation of risks of medical side effects, for which the assigned frequencies centred around small numerical values (Berry et al., 2002). Although the range of the scales in these contexts reflect real-world distributions (e.g., a food would typically not contribute more than 30% of one's GDA of a nutrient; Rayner et al., 2004), if people do not realise what the guideline ranges encompass, they will misinterpret the verbal terms.

The extent of misalignment with the recommended ranges was substantial: in the student sample, less than a quarter of interpretations were within the recommended range. Although the nationwide sample performed better, still around half the interpretations exceeded guidelines. Beyond the proportion of people overestimating, the extent of the overestimation suggests a further problem with the guideline ranges, especially for 'high'. These ranges use an official cut-off (e.g., $> 30\%$ for minerals), which offers substantial vagueness in meaning. High % was commonly translated as around 68% in Experiment 1 and 48% in Experiment 2, which could indicate an overestimate of where the cut-off actually lies, even if technically the estimate is correctly within the $> 30\%$ bracket. People's large overestimation of 'high %' also indicates that the guideline may be too vague to be meaningful, as it is misaligned with people's natural interpretations of the word. Misinterpreting a lower boundary would have serious implications for people's understanding about their diet quality, as people could believe they are consuming more of these nutrients than they actually are. This would be primarily a problem for under-consumption of food groups like vitamins, minerals, and fibre. The belief that one's own diet already meets dietary recommendations for these food groups can prevent one from taking action to improve dietary behaviour (Lechner et al., 1998). For example, manufacturers can declare a food high in fibre if it reaches 30% of the GDA. However, if one believes it provides 70% of a day's recommended fibre consumption, one would mistakenly assume that one has almost reached their fibre target for the day, and not seek to improve their fibre intake. Fibre consumption in the general population is far below recommended amounts (Guiné, R. P. F., Duarte, J., Ferreira, M., Correia, P., Leal, M., Rumbak, I., ... Straumite, E., 2016); 97% of adults in our sample population already do not meet daily fibre targets (NatCen Social Research, 2015). Both overconsumption of negative nutrients and under-consumption of positive ones impact health (UK Department of Health, 2015), so it is important for future work to determine whether, and to what extent overestimating the nutrient content of one's food affects subsequent consumption decisions.

A further problem about misinterpreting verbal quantities arises because information about different nutrients is often presented with different formats on the same product. When comparing between verbal and numerical quantifiers, people may reach incorrect conclusions about the levels of nutrients within one product. For example, a breakfast cereal that is 'high in fibre' while providing a 30% GDA of sugar might seem to have a lot more fibre than sugar when in fact both nutrients contribute equally to their respective GDAs. Finally, differences in interpretation of verbal and numerical labels can result in erroneous comparisons between products if one uses a verbal quantifier and the other a numerical quantifier. For example, one might mistake a cereal with 12% GDA of fat as healthier than a low-fat cereal if one believes 'low' to indicate 20%. Although past work indicates that consumers can use labels in the same format to pick between products with higher or lower nutrient contents (Hersey et al., 2013), it is not known how they would perform when comparing between foods with different label formats. Our work is a first step to demonstrate how discrepancies in interpretation could potentially affect food choice. A question for future research is to investigate whether subsequent judgements and decisions about food are indeed affected by interpretational variance of verbal quantifiers.

2.5.3 Nutrient valence, but not personal characteristics, affect interpretations

We found in Experiment 2 that people perceived quantities and produced numerical translations of minerals (a positive nutrient) as larger than quantities of fat (a negative nutrient), although this nutrient valence effect was not present when participants saw many nutrients one after another (Experiment 1). In line with a motivated reasoning account of quantifier interpretation (Kunda, 1990), participants might wish to believe that they were not consuming too much of a negative nutrient, but were eating more of a positive one. However, we did not find evidence that individual differences in motivation to eat healthily affected participants' interpretations, nor did any other individual difference variable predict participants' estimations. This could reflect that individual predictors for making healthier choices (e.g., motivation to eat healthily; Hearty et al., 2007) or better health literacy (e.g., education level: Sinclair et al., 2013; familiarity with labels: Gigerenzer et al., 2005) do not guard against overestimation of verbal quantifiers. Our findings contrast with studies in verbal probabilities, where differences in attitudes significantly predicted interpretational variability (Budescu et al., 2014), and studies in verbal frequencies, where gender, ethnicity, and education levels predicted overestimation of risk frequencies (Webster et al., 2017). This could be because attitudes are less predictive of quantity perception for food than for contexts such as climate change, where one's attitude is more closely related to their beliefs (Hornsey et al., 2016). More frequent label use may also not offer the right feedback to inform people that they are overestimating verbal quantities. Therefore, the individual difference variables we measured may not have fully captured the factors that might explain variance in interpretations. Beliefs about specific nutrients and their contributions to health (e.g., Oakes, 2005a) might better explain tendencies to overestimate quantifiers more than general attitudes towards eating healthily. It would be good to investigate further if people's cognitions about different categories of nutrients can form a stable factor to determine their tendency to overestimate a quantity.

2.5.4 Improving interpretation accuracy: Challenges for food labelling policy

Interestingly, participants who frequently consulted nutrition labels were no less likely to overestimate verbal quantifiers. The GDA system is the most common front-of-package label in the UK (Storcksdieck genannt Bonsmann et al., 2010), so we could assume that frequent label checkers are familiar with this information. However, exposure to labels may not confer understanding (Grunert & Wills, 2007). Indeed, consumers must independently look up advice on how to interpret verbal quantifiers in terms of numerical percentages if they wish to ensure an accurate interpretation. Alternatively, consumers could consult detailed nutrition information that specifies exact numerical amounts of each nutrient to perform the necessary calculations and confirm the quantities. This information is typically relegated to the back of food packaging (Storcksdieck genannt Bonsmann et al., 2010). Given that even health-conscious consumers do not fully check back-of-pack labels for detailed information (Hieke & Taylor, 2012; Higginson et al., 2002), the average consumer is unlikely to be motivated to perform this detective work to understand which quantities should be considered 'low', 'medium', or 'high'. Based on our experimental findings, it is likely that they will simply calibrate verbal quantifiers to a full percentage scale.

A potential solution might be to standardise GDA labelling across all food products to include both verbal and numerical information (e.g., the hybrid Traffic

Light system; Limb, 2012). Dual (verbal-numerical) scales are intended to increase interpretational consistency (Budescu et al., 2012), however there is no guarantee that people will not rely on one format or the other, especially if it is unclear how the scale values are derived (Nicolas et al., 2010). Greater reliance on text over numerical information is also seen in people with low numeracy (Dieckmann et al., 2009), who tend to prefer non-numeric information (Peters, 2012; Reyna et al., 2009). Nevertheless, the dual-scale labels may pose an advantage over number-only labels, which are more likely to be ignored due to the higher cognitive effort involved in processing them (Cowburn & Stockley, 2005).

Alternatively, research from the medical domain advocate the use of graphic representations of quantitative information, for instance, using discrete frequencies to represent probabilities (e.g., icon arrays; Zipkin, D. A., Umscheid, C. A., Keating, N. L., Allen, E., Aung, K., Beyth, R., ...Feldstein, D. A., 2014). Such formats facilitate understanding of health risks, particularly among less numerate or literate segments of the population (Gigerenzer & Kolpatzik, 2017). In practice, consumers often make food choice decisions quickly (Celnik et al., 2012), and food packaging has limited space available to provide information. Alternative formats attempt to address these issues by interpreting the food product for consumers (Maubach et al., 2014). For example, a label may classify products as either healthy or not (e.g., tick logos: Scott & Worsley, 1994; the Smart Choices label: Roberto et al., 2012), give a product a healthiness score (e.g., the Star rating; Maubach et al., 2014), or use colour coding to draw attention to the level of a nutrient (e.g., Traffic Light systems; Mejean et al., 2013).

Interpretive labels could help people understand labels by tapping into their tendency to form categorical representations of quantitative information. For instance, they might form a rough summary of the information, such as ‘the food has fibre’ or ‘the food has fat’ to make their decision (Blalock et al., 2016).However, food information must still align with consumer expectations if they are to provide an accurate picture of the quality of one’s diet (Celnik et al., 2012). For example, the colour bands in the Traffic Light system (red, amber, and green) correspond to verbal bands of high, medium, and low (UK Department of Health, 2016). Even if colour-coded labels are easier to process (Siegrist et al., 2015), consumer misinterpretations of the low, medium, and high bands could still lead to mistaken assumptions about the quality of their diets. Thus, the issue of whether verbal bands (or other interpretive text) match psychologically equivalent numerical values remains important. Current guideline ranges for determining verbal banding were developed based on scientific research on nutritional values suitable for populations, and allow for discrimination between the low-skewed GDA percentages typically seen on individual foods (Rayner et al., 2004). However, guidelines for verbal labels are determined by GDA percentages for some, but not all nutrients (UK Department of Health, 2011, 2016). This increases the likelihood that people will revert to natural interpretations when reading nutrition labels. These interpretations appear to be independent of the specifications of food science, show little discrimination between small numerical differences, and are difficult to override. Indeed, research shows that even when given explicit interpretational guidelines to consult, people nonetheless provide numerical estimates for verbal quantifiers that do not conform to guidelines (Budescu et al., 2012; Webster et al., 2017).

The use of verbal labels may simplify nutritional information and nudge consumers towards healthier food choices (Thaler & Sunstein, 2008). Combined with public education about what nutritional levels the verbal bands indicate, it is hoped that people will also have accurate knowledge about their personal nutrition. However, we provide strong evidence that governmental efforts to educate the public about verbal banding on nutrition labels (UK Food Standards Agency, 2007, 2008) did not enable our participants to suppress their natural interpretations of verbal and numerical quantifiers. Should guideline ranges therefore be refined to take into account people's natural interpretations of verbal quantifiers? Given our study's findings, we believe this question needs to be addressed. Perhaps it would be beneficial to supplement efforts to improve consumers' nutritional knowledge with a recalibration of the ranges assigned to verbal quantifiers. This might strike a balance between the ability to easily distinguish between different nutrient quantities and boosting nutritional understanding (Hertwig & Grüne-Yanoff, 2017) by harnessing the intuitive understanding of what these quantifiers mean. Our findings regarding the great inter-individual variability in interpretations suggest that identifying the best interpretive range for verbal nutrient quantifiers would be a challenge in practice. Nonetheless, we suggest two immediate concerns with current guidelines that could be addressed. First, guidelines on the use of verbal quantifiers such as ‘high’ for positive nutrients could be pegged higher to avoid problematic overestimation of positive nutrient consumption. Second, interpretational guidelines could be standardised across nutrients and products to facilitate comparisons.

2.5.5 Conclusion

Legislation to tackle obesity must strike a balance between paternalistic interventions (e.g., sugar taxation; HM Revenue and Customs, 2016) and cognitive strategies such as nutrition labelling (Malam et al., 2009). Nutritional information such as GDAs empower consumers with the ability to make informed choices as long as they correctly understand the nutritional value of the food they are eating. However, we showed that the interpretation of quantifiers in verbal and numerical formats is variable across individuals and misaligned with recommended guidelines, resulting in misinterpretations that can affect the conclusions people draw about food and their daily consumption levels. More work needs to be done to identify the best verbal ranges to attach to numerical quantities in official communications so that people will interpret them as intended. We hope that our work can open this debate and stimulate more empirical research to support continued refinement of nutritional guidelines that better suit people’s intuitive understanding of quantifiers.

Chapter 3: Differences Between Decisions Made Using Verbal or Numerical Quantifiers

3.1 Abstract

Past research suggests that people process verbal quantifiers differently from numerical ones, but this suggestion has yet to be formally tested. Drawing from traditional correlates of dual-process theories, we investigated whether people process verbal quantifiers faster, less accurately, and with less subjective effort than numerical quantifiers. In two pre-registered experiments, participants decided whether a quantity (either verbal or numerical) of a nutrient, summed with a pictorial quantity, exceeded a recommended total. The verbal quantifiers were matched to average numerical translations (Experiment 1) as well as translations from participants themselves (Experiment 2). Across experiments, participants did not answer faster or find verbal quantifiers less effortful than numerical ones, but they made less accurate decisions on average with verbal quantifiers because they used more context-based decision shortcuts (e.g., 'minerals are healthy'). Our findings suggest that it is how much people rely on context that distinguishes their decisions with verbal and numerical quantifiers.

3.2 Introduction

The study of decision-making commonly investigates choices made between options that require comparisons or evaluations, both of which regularly include quantities. For instance, which option has a greater chance of success, or offers the best value for money? These quantities can be expressed using quantifiers of different formats: either numerical or verbal. For example, a person might decide that a food is healthy because it provides 30% of their recommended daily amount of fibre, or simply if it is 'high' in fibre. The way people make decisions involving verbal and numerical quantifiers suggests that people process the meaning of these quantifiers differently, with numbers requiring more effort to process than words (Childers & Viswanathan, 2000). This view that numerical formats are more effortful to process than verbal ones drives the use of verbal formats to communicate quantities in many different applied contexts (e.g., nutrition information, Malam et al., 2009; medical risks, Berry et al., 2003). One suggested explanation for the processing difference is that people adopt a more intuitive approach to words and a more analytical one to numbers (Windschitl & Wells, 1996). However, there remains a paucity of empirical research that directly compares how the format of a quantifier affects processing style. Therefore, in the present research, we addressed this gap by testing the cognitive processing styles for verbal and numerical quantifier formats.

3.2.1 Processing differences in verbal and numerical quantifiers

Evidence from non-comparative studies that have investigated in isolation either verbal or numerical quantifiers support different conclusions about how people process the two formats. People process the meanings of words automatically, as illustrated by the Stroop task, where people faced a conflict in naming the colour of a word written in, for example, yellow font, if the word spelt 'blue' (MacLeod, 1991). Studies of verbal probabilities show similar conflicts between what people automatically understand verbal quantifiers to mean, and what they officially mean: people were given official verbal-numerical descriptors for how likely a climate event would occur, but continued to provide translations for the verbal probabilities that did not match the stated guidelines (Budescu et al., 2012). In addition to giving an automatic sense of the magnitude of an amount, verbal quantifiers may also help to contextualise this amount and help people understand the focus of the information. For example, verbal probabilities can either focus on the chance of an event occurring (e.g., 'it is likely to happen') or not occurring (e.g., 'it is unlikely to happen'; Teigen & Brun, 1995, 2000, 2003; Teigen et al., 2014). This may aid evaluations made with verbal quantifiers: for instance, people could determine without much effort where a cereal with 'high protein' lies on an evaluative scale (i.e., good or bad), and quickly judge that product (Viswanathan & Childers, 1996). Finally, people have reported that verbal expressions of food quantities were easier to use in decision-making (Malam et al., 2009).

On the other hand, people must often expend deliberate effort to process numbers. Numerical calculations require use of working memory (DeStefano & LeFevre, 2004), and numerical probabilities (e.g., 30% chance) are less clear as to whether they refer to the possibility of an event, or the possibility that it will not happen (Teigen & Brun, 1995, 2000). This could be why understanding where numbers lie in an evaluative context (e.g., whether 20% protein is good or bad) takes more time (Viswanathan & Childers, 1997). Finally, people have reported that numerical quantifiers are cognitively effortful and difficult to understand (Peters et al., 2007). Altogether, these findings present a picture of verbal quantifiers that is more automatic and effortless, and a picture of numerical quantifiers that is more deliberate and effortful.

When verbal and numerical formats were directly compared on measures that should distinguish processing style differences, evidence appears to be inconclusive. If numerical quantifiers require more effortful processing, people should need more time to complete tasks with numerical than verbal quantifiers. This was seen when a task required participants to understand the meaning of quantifiers (e.g., assessing its position on a scale). In this case, people were quicker to make judgements for verbal than numerical quantifiers (Childers & Viswanathan, 2000). However, in tasks where participants had to compare the magnitudes of two numerical or two verbal quantifiers, people were quicker with the numerical quantifiers (Jaffe-Katz et al., 1989; Viswanathan & Narayanan, 1994). More recently, Shikhare et al. (2015) investigated how long participants took to judge if numerical and verbal quantifiers descriptions of visual displays were correct (e.g., 'at least seven balls are yellow' vs. 'many balls are yellow'). Although mean response times for numerical quantifiers were higher than those for verbal quantifiers, the researchers analysed the formats separately and did not directly compare them.

Despite mixed evidence about processing speeds for one quantifier format over another, findings that verbal and numerical formats lead to different decision outcomes may still indicate different cognitive processes. Although people's aggregated response times and decision performance did not differ between verbal and numerical formats, people would select gambles with verbal probabilities (e.g., 'likely', 'probably') more when the gambles paid better; in contrast, they would select gambles with numerical probabilities (e.g., 'a 60% chance') more when the gamble was more likely to succeed (González-Vallejo et al., 1994). This suggests that participants used different pieces of information to reach the final decision: they relied more on contextual aspects of a problem when given verbal quantifiers, i.e., how positive the outcome would be.

3.2.2 Processing quantifiers: A dual-process perspective

The differences associated with verbal and numerical quantifier processing appear to fit dual-process theories about the human mind (De Neys, 2017b; Evans & Stanovich, 2013; Kahneman, 2011; Sloman, 1996). The generic theory, of which there are a number of variants (Pennycook et al., 2018), describes two processing styles that differ in terms of consciousness, automaticity, and the amount of cognitive effort involved (Evans, 2008). Within this framework, the processing of verbal quantifiers could be viewed as more intuitive: automatic, unconscious, and quick, generating affective cognitions such as a feeling of rightness about the decision. On the other hand, numerical quantifiers could require more analytical processing: requiring conscious, effortful processing, which operates more slowly and deliberately (for an overview of dual-process theories, see Evans, 2008, and Evans & Stanovich, 2013; for alternative and more critical views of intuitive and analytical processing, see Betsch & Glöckner, 2010, and Melnikoff & Bargh, 2018). Verbal and numerical quantifiers also seem to suit intuitive and analytical tasks respectively. People prefer verbal quantifiers to numerical ones with tasks that required subjective judgements and relied on affective information (Nicolas et al., 2010; Wallsten et al., 1993), or involved preference-based judgements (Wilson & Schooler, 1991). Such features are also often associated with intuitive tasks (Hammond, 1988). This suggests that intuitive decision-making may suit the way people naturally process verbal quantifiers. Conversely, people prefer numerical quantifiers with tasks that involve analytical judgements and objective values (Budescu & Wallsten, 1990), suggesting that analytical processing is more suitable for numerical quantifiers.

The traditional correlates of intuitive (vs. analytical) processes in dual-process theory are that intuitive processes should operate quicker and require less effort. Although there is mixed evidence comparing decision speed and performance for verbal and numerical quantifiers (Childers & Viswanathan, 2000; González-Vallejo et al., 1994; Jaffe-Katz et al., 1989; Viswanathan & Narayanan, 1994), this could reflect differences in the nature of the tasks used. People were quicker with verbal than numerical quantifiers when the task required evaluation (Budescu & Wallsten, 1990; Viswanathan & Childers, 1996), but quicker with numerical than verbal quantifiers when the task involved comparing simple magnitudes (Viswanathan & Narayanan, 1994). Hence, one could expect that in the context of decision-making encompassing a mix of evaluation and calculation processes, verbal quantifiers should be quicker and less effortful —both indicators of a more intuitive processing style.

Intuition is also typically associated with the use of mental shortcuts to avoid the higher cognitive demands of analytical reasoning, often resulting in decision biases and errors (Kahneman, 2011). For example, people may rely on contextual knowledge to substitute a more mentally available concept to answer a difficult question (Kahneman & Frederick, 2002). People often use contextual knowledge to guide their interpretation of quantified sentences (Dwivedi et al., 2018), and some evidence suggests that verbal quantifiers are more susceptible to the influence of contextual information (González-Vallejo et al., 1994; Windschitl & Wells, 1996). For instance, people's gambling decisions were more correlated with payoff outcomes for verbal than numerical quantifiers, suggesting that this information, which contextualises the probability of winning, was more important to the decisions with the verbal quantifiers (González-Vallejo et al., 1994). However, when dealing with gambles, it is possible that numerical processing was also engaged in a verbal quantifier decision, since the outcome (a payoff) was also presented as a number. Further work demonstrating intuitive biases with verbal but not numerical quantifiers also used numerical formats for the contextual information, asking participants to generate verbal or numerical probability estimates to describe two equivalent numerical frequencies (e.g., describing a '1 in 10' or a '10 in 100' chance; Windschitl & Wells, 1996). In such a case, it is uncertain whether the finding that participants produced different verbal but not numerical probabilities for the two frequencies was a product of incompatibility between the format of the contextual information and the response (numerical-numerical vs. numerical-verbal) rather than different processing approaches to the quantity information. Therefore, to test the hypothesis that verbal quantifiers result in poorer performance and more reliance on the context, an investigation needs to use a task where only the quantifier, and not the contextual information, can take numerical format.

3.2.3 The present work

Our investigation aimed to test for the first time, the direct effect of format on the processing of quantifiers by focusing on three indicators of processing styles while overcoming methodological issues mentioned above. Focusing on dual-process theory as a framework, we investigated several traditional correlates of intuition and analysis as indicators of processing style (Evans & Stanovich, 2013). Based on a dual-process classification, intuitive processes are expected to produce decisions that are quicker and less effortful than analytical processes (Evans, 2008). However, intuitive processes are also expected to rely more on mental shortcuts (e.g., contextual information) that may hinder decision-making performance (Kahneman, 2011). We therefore expected that participants would make decisions with verbal quantifiers quicker (Experiments 1 & 2) and with less effort (Experiment 1) than numerical quantifiers. We also expected that participants would make less accurate decisions with verbal quantifiers than numerical ones (Experiments 1 & 2), because they relied on the contextual information more (Experiment 2).

In addition, we extended previous research on quantifiers to a novel context, nutrition communication. We chose this context because it fulfils three important criteria. First, it allowed us to design a task that was less geared towards numerical processing, as the context of previous work has been (e.g., calculating gambles and probabilities; González-Vallejo et al., 1994; Windschitl & Wells, 1996). This meant that it should not already trigger a more analytical style based on the information context. Second, nutrient quantities are commonly expressed using verbal as well as numerical formats (Malam et al., 2009). As such, there is precedent for both formats being used in the real world. Finally, the nutrients provide contextual information that can be positive or negative (e.g., minerals vs. sugar). This allowed us to easily manipulate the valence of the contextual attributes. Based on these criteria, we designed a quantity integration task that required participants to make a decision based on either a verbal or numerical nutrient quantifier, with previous information presented in a pictorial format to minimise priming of verbal or numerical quantifier processing.

3.2.4 Open science statement

The hypotheses, methods, and analytical strategies were registered prior to data collection. Both pre-registrations, along with the materials and data, are available on the Open Science Framework (OSF).

3.3 Experiment 1

3.3.1 Method

Participants. The study was powered to detect a small-to-medium effect in a mixed ANOVA (Cohen's f = .18, α= .05) with 80% statistical power. Ninety-three participants were sourced from a university lab database and paid £8 for their participation (67% female; age range: 18-67, M = 22.37, SD = 6.76). All participants had completed at least high school education, and 47% also had a university degree. Participants' racial background was 47% White, 37% Asian, and 11% African.

To control for the usual processing styles of participants and their attitudes towards food, we measured at the end of the experiment participants' preferences for intuition and deliberation (Betsch, 2004), eating attitudes (Steptoe et al., 1995), and BMI (derived from height and weight). Our sample displayed a slight preference for deliberation over intuition (M_{diff} = 0.26, SD_{diff} = 0.71). They reported a positive attitude towards healthy eating (M = 5.11, SD = 1.09). Mean estimated BMI was 22.56 (SD = 4.39; this is in the healthy range), and 51% reported general use of nutrition labels in everyday life.

Design. Participants decided whether it was healthy to eat a given quantity of a nutrient. The quantifier was either verbal or numerical (manipulated between-subjects). We aimed to test the effect of format on decision-making in a range of decisions, hence we manipulated three other variables within-subjects: the type of nutrient (fat, sugar, and minerals), the quantity (very low to very high; see Table 1), and the correct decision (whether the quantity was within or exceeded limits). We therefore employed a 2 (format) $\times$ 3 (nutrient) $\times$ 5 (quantity) $\times$ 2 (correct decision) mixed design.

Materials. We created a decision task programmed using Inquisit 4 (Millisecond Software, 2015; code available on the OSF). In this task, participants decided whether a given standardised percentage of a nutrient could be eaten without exceeding their guideline daily amount (GDA). This is the most common presentation of nutrition information in the UK and the EU (Storcksdieck genannt Bonsmann et al., 2010). The task instructions explained the concept of GDA, and specified how to decide whether it was healthy to consume a quantity based on its GDA value, as follows:

> *Your Guideline Daily Amount (GDA) is the* ***total amount of a nutrient that you should consume in a day*** *as part of a healthy diet. GDA labels on food tell you the contribution of the food towards your GDA* ***for that nutrient*** *as a percentage. This will help you to decide if it is healthy to consume a food based on how much it adds to your daily recommended total. For example, a GDA of 25% for fat means that the food gives you 25% of the fat you should eat in a day.*

Because GDAs are usually calculated based on dietary requirements for a typical person (Rayner et al., 2004), we also included an instruction that participants should assume the GDAs in the task were tailored to their own dietary needs. This was to control for variations in how much people might view the concept as applicable to them.

Participants then saw a pie chart and food label example, with instructions on how to complete the task, as illustrated in Figure 3.1. These instructions specified that participants should decide if consuming the amount on the label would be healthy in the context of what they had already consumed of their GDA. The task was comprised of 30 decision trials formed from variations of the within-subjects conditions, presented in a randomised order. Each decision trial had two components, as depicted in Figure 3.1.

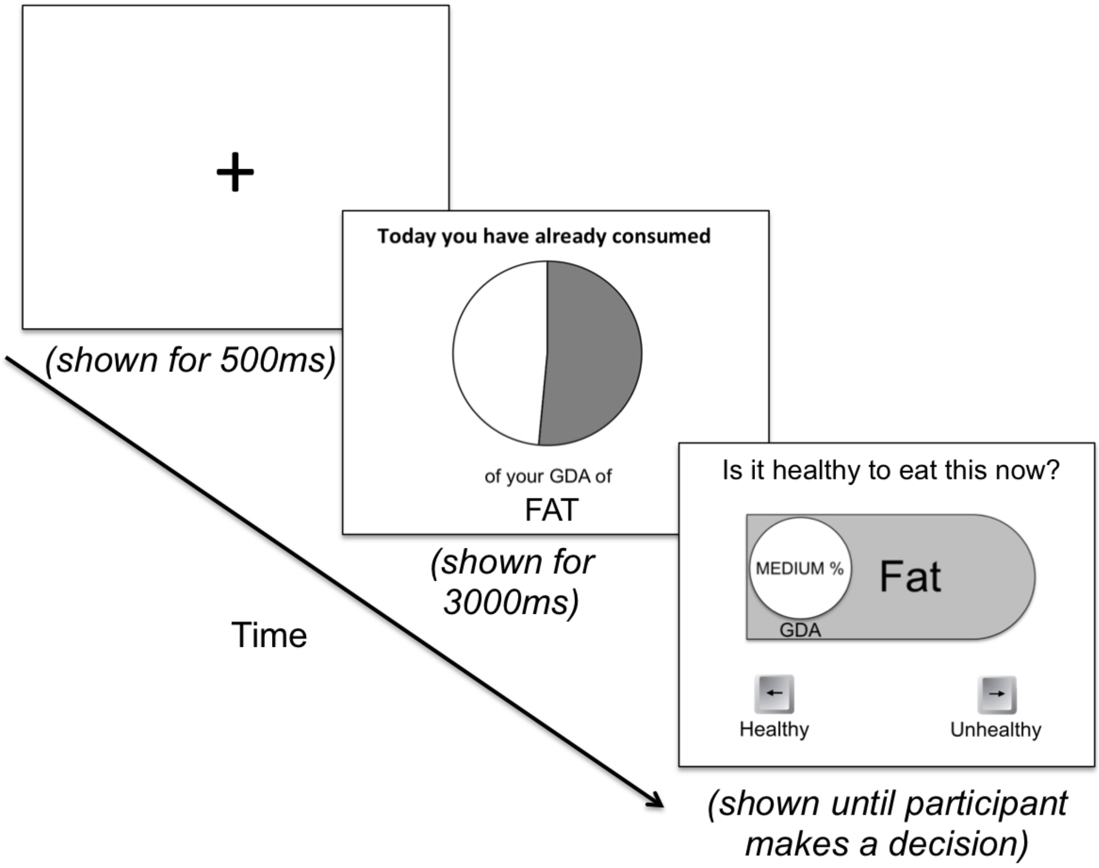

Figure 3.1. GDA decision task and instructions received for the task. The three consecutive screens constituted one trial. Instructions were given with illustrations of the task stimuli (pie chart and label screens) at the start of the experiment.

Communicating prior consumption. In each decision trial, participants imagined they had consumed a given quantity of a nutrient, shown as the shaded area in a pie chart. We used a pie chart to present this information so as not to prime participants with either a verbal or numerical quantifier prior to the main decision that we were measuring. Combining a pictorial quantity with the label quantity (verbal or numerical) also meant that we could compare decisions with the verbal and numerical quantifiers without the confound of one quantifier format matching that of the first-presented quantity. The pictorial quantity gave a level of precision between the precise numerical format and the vague verbal one. We considered this appropriate because it was vague enough to prevent simple addition with the numerical quantifiers, but precise enough to allow participants to add the vague verbal quantifiers.

Communicating nutrient quantities. Nutrient quantities on the food labels were presented with either a verbal or a numerical quantifier. Following methods in other studies comparing verbal probabilities with their average numerical translations (Teigen & Brun, 2000; Welkenhuysen et al., 2001), we matched numerical quantifiers to verbal ones (columns 3 and 4 of Table 3.1) using established translations for verbal expressions of nutrient quantities (see Chapter 2). We used the average translations found in the first study of Chapter 2 as these had been found for a similar sample. We did not rely on guidelines on how to translate verbal quantifiers in the food industry because there is substantial evidence that people do not perceive the magnitude of verbal quantifiers in line with existing guidelines (Berry, 2006; Berry et al., 2002; Budescu et al., 2014; Knapp et al., 2009a, 2010).

Decision. Participants decided, according to the rules of the task, whether consuming this amount would fall within their GDA limit. Because the task was a mathematical one (typically analytical), we sought a way to allow a more intuitive response. This gave us a better chance to deduce whether participants relied more on intuitive processes. We therefore set the judgement keys as 'healthy' (i.e., the quantity was within limits) or 'unhealthy' (i.e., the quantity exceeded limits). The healthy button was either the left or the right arrow key. One could therefore give an intuitive answer based on intuitions about whether a nutrient was healthy, whereas an analytical answer would require participants to perform the calculation steps of comparing the quantities and integrating it with the guideline definition given in the instructions. Participants practised the decisions prior to starting the trials, with feedback given to explain why a within-limits combination was healthy, and an exceeded-limits combination was unhealthy.

Measuring response time. We measured how long participants took to make their decision. Although we did not pre-register data transformations prior to analysis, we found that the distribution of response times had substantial positive skew (skewness = 2.28, 95% CI [2.13, 2.42]). Hence, we log-transformed response times prior to analysis. We also excluded 5 trials (0.2%) where the response time was below the threshold for manual response to a visually-perceived stimulus (< 150ms; Amano et al., 2006) and 39 response times (1.4%) that exceeded 10s as these latencies (> 5 *SD* above the mean) suggested that participants were not responding immediately to these trials.

Measuring decision quality. Decision-making performance was determined by whether participants correctly identified the quantities as fitting the limit (i.e., healthy to consume) or exceeding the limit (i.e., unhealthy to consume). Each quantity was combined with one pie chart that would be within limits —'healthy', and one pie chart that would exceed limits —'unhealthy'. The magnitude of the quantity shown in the pie chart was derived from the average and standard deviation of the average numerical meaning of the five studied verbal quantifiers (see Table 3.1), as measured in Experiment 1 of Chapter 2, which used a similar sample. The healthy (within limits) pie chart magnitude was computed as 100% - ($M_{verbal\ quantifier\ translation}$ + 1 $SD_{verbal\ quantifier\ translation}$) and the unhealthy (exceeding limits) pie chart magnitude was computed as 100% - ($M_{verbal\ quantifier\ translation}$ - 1 $SD_{verbal\ quantifier\ translation}$). In the cases where this rule resulted in pie chart values above 99% or below 1%, '1 *SD*' was replaced by '0.5 *SD*' in the formula. For example, for 'low %', the pie chart within limits was 66.98% (20% + 66.98% < 100%, thus the combination is healthy), and the pie chart exceeding limits was 91.13% (20% + 91.13% > 100%, thus the combination is unhealthy).

Measuring subjective effort. After every fifth decision trial, participants reported how cognitively effortful they found the task by clicking on a 5-point Likert scale (anchored as 1: *very hard*, 5: *very easy*).

Procedure. After giving informed consent and reading the instructions, participants performed a training decision block before the experimental phase. In the practice phase, participants received performance feedback. If they incorrectly judged a within-limits quantity as unhealthy, they were informed: '*Your GDA is*

Table 3.1. Quantity combinations and their correct decision in the task trials. These 10 combinations were repeated across three nutrients (fat, sugar, and minerals) to create 30 decision trials.

Pie chart value	**Quantity**		**Correct decision**
	Verbal	**Numerical**	
78.29%	Very Low %	10 %	Within limits (healthy)
66.88%	Low %	20 %	Within limits (healthy)
44.79%	Medium %	40 %	Within limits (healthy)
12.22%	High %	70 %	Within limits (healthy)
11.51%	Very High %	80 %	Within limits (healthy)
96.82%	Very Low %	10 %	Exceeds limits (unhealthy)
91.13%	Low %	20 %	Exceeds limits (unhealthy)
68.95%	Medium %	40 %	Exceeds limits (unhealthy)
51.46%	High %	70 %	Exceeds limits (unhealthy)
42.26%	Very High %	80 %	Exceeds limits (unhealthy)

Note. We obtained the numerical quantifiers for the study and derived a previous quantity that would fall within or exceed limits for each based on the distributions of verbal-numerical translations in 2.

the total amount of the nutrient you can eat in one day. By eating this food, you would not exceed this total. It is healthy to stay within the recommended total for the day.' If they incorrectly judged an exceeded-limits quantity as healthy, they were informed: '*Your GDA is the total amount of the nutrient you can eat in one day. By eating this food, you would exceed this total. It is not healthy to eat an amount that will cause you to exceed your recommended total for the day.*' If they were correct, they received a similar explanation for why they were correct. Participants could not proceed to the experimental phase until they had performed the final of three practice trials correctly.

In the experimental phase, participants received no feedback on their performance. The experimental phase had six blocks of five decision trials. At the end of each block, participants completed the effort measure and were offered a break before continuing. The experiment also included an additional six blocks where participants were instructed to be either intuitive or analytical with their decisions (Schroyens et al., 2003). This was intended to test an additional pre-registered hypothesis that participants' natural decisions with verbal quantifiers would match decisions made when told to be intuitive, and participants' natural decisions with numerical quantifiers would match decisions made when told to be analytical. We ran a manipulation check at the end of the instructed blocks, which asked participants to report how they completed the task in relation to 10 adjective pairs that described intuition on one end and analysis on the other (e.g., *quickly –slowly*, *automatically –systematically*). This manipulation check revealed that the instruction participants received had no significant difference in self-rated approach to decision-making in the task, $t(91) = 1.55$, $p = .125$. We also found that instruction type had no effect on the dependent variables, $F(3, 89) = 1.55$, $p = .206$, $\eta^2_P = .05$ (using Pillai's trace in a MANOVA testing instruction type as a factor). We have thus not included these trials in the main analysis of the data, and do not report further these results in this manuscript. However, these analyses are included in Appendix B. Data from these trials is also archived on the OSF.

3.3.2 Results

Mean response time, performance, and effort were not significantly different between formats. For each participant, we calculated an average response time, task performance, and effort rating across all the experimental trials (see Table 3.2). On average, the numerical formats showed a trend for better performance, quicker decisions, and more effort required than verbal ones. However, the pre-registered multivariate analysis of variance testing for the effect of format on mean response time, task performance, and effort ratings showed no statistical significance effect of format, $F(3, 89) = 1.03$, $p = .384$, $\eta^2_P = .03$ (all tests for format effects at each individual dependent variable were also non-significant, > .222).

Table 3.2. Descriptive summary of response times (ms), decision performance (% of correct trials), and effort (rating from 1: very hard to 5: very easy) across experiments and quantifier formats.

	Experiment 1		**Experiment 2**	
	Verbal	**Numerical**	**Verbal**	**Numerical**
Response time (untransformed)				
Median	953ms	904ms	1711ms	1654ms
Inter-quartile range	291ms	326ms	855ms	928ms
Response time (log-transformed)				
Mean (*SD*)	2.96 (0.18)	2.94 (0.15)	3.20 (0.21)	3.18 (0.20)
95% CI	[2.91, 3.01]	[2.90, 2.99]	[3.16, 3.25]	[3.13, 3.23]
Performance (% of correct trials)				
Mean (*SD*)	70.82 (11.70)	74.72 (18.50)	61.71 (21.03)	78.13 (17.62)
95% CI	[67.46, 74.18]	[69.10, 80.35]	[57.31, 66.12]	[73.57, 82.68]
Effort (rating on 5-point scale)				
Mean (*SD*)	3.59 (0.78)	3.76 (0.67)		
95% CI	[3.36, 3.81]	[3.55, 3.96]		

Exploratory analyses. Although on average (i.e., across trials), participants did not differ significantly according to the quantifier format, we observed that mean performance varied across the nutrient types, quantities, and correct decision conditions. In our task, there were two decision situations that were counter-intuitive based on the correct decision and the nutrient involved: identifying that a positive nutrient quantity was unhealthy (independent variables: nutrient —minerals and correct decision —healthy) or that a negative nutrient quantity was healthy (independent variables: nutrient —fat or sugar and correct decision —unhealthy; illustrated in Figure 3.2). As shown in Table 3.3, participants made more errors with verbal than numerical quantifiers when the correct decision (within or exceeding the limit) for minerals was unhealthy (exceeding limits), thus conflicting with the nutrient's valence (positive). This suggested that based on verbal quantifiers, participants relied more on the valence of the nutrient (although it was irrelevant) to reach their decision, instead of only focusing on the quantities themselves.

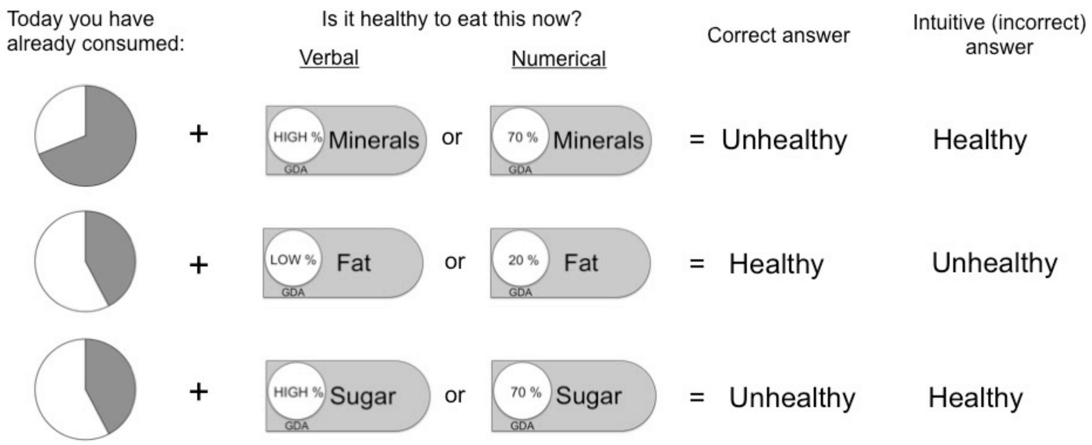

Figure 3.2. Examples of nutrient and quantity combinations that had an incorrect intuitive answer.

To assess whether format affected participants' use of the contextual information (e.g., the type of nutrient) to make the decision, we opted for an exploratory analysis using a multilevel model. This approach allowed us to examine all trials in the long-form data (rather than aggregating responses across nutrient, quantity, and correct decision). We were therefore able to test the effect of format, together with nutrient, quantity, and correct decision, along with their interactions, on performance. The model used a variance components matrix and included random by-participant intercepts to account for individual variations among participants (the full random effects model, factoring in individual responses to the fixed factors, failed to converge, thus we removed random slopes

Table 3.3. Proportion of errors made in decision task with verbal and numerical formats when the normative response conflicted or did not conflict with an intuitive response.

	Proportion of errors					
	Verbal			Numerical		
	Conflict	**No conflict**	**95% CI of M_{diff}**	**Conflict**	**No conflict**	**95% CI of M_{diff}**
Experiment 1						
Fat	22.08%	28.94%	[-1.17, 14.90]	20.01%	15.17%	[-2.52, 12.18]
Sugar	23.01%	26.80%	[-4.22, 11.80]	26.13%	19.72%	[-1.93, 14.76]
Minerals	51.40%	11.88%	[31.14, 47.90]	26.51%	19.67%	[-1.58, 15.25]
Average	32.16%	22.54%		24.22%	18.19%	
Experiment 2						
Fat	41.51%	29.57%	[0.15, 23.74]	19.34%	22.01%	[-8.96, 14.31]
Minerals	65.12%	20.63%	[34.10, 54.88]	50.95%	4.57%	[34.36, 58.41]
Average	53.32%	25.10%		35.15%	13.29%	

Note. Conflicts were trials on which participants were given healthy combinations of fat and sugar or unhealthy combinations of minerals (as illustrated in Figure 3.2).

until we achieved a convergent model). We included format, nutrient, quantity, and correct decision, and their two- and three-way interactions as fixed factors. In particular, to test whether participants made more errors in deciding whether a quantity was healthy by relying on the nutrient with verbal than numerical quantifiers, we were interested in the interaction of format, nutrient, and correct decision, and the pairwise comparisons between each of these factors. The results of the multilevel analysis are summarised in Table 3.4.

Participants used more context-based shortcuts with verbal quantifiers. The significant two- and three-way interactions between format, nutrient, and correct decision are illustrated in Figure 3.3, and showed that decision performance varied across combinations of these factors. Participants made more errors when the correct decision was counter-intuitive (e.g., an unhealthy quantity of minerals). These intuitive errors were on average more common for

Table 3.4. Results of the multilevel analysis on decision performance in Experiments 1 and 2 (effects specifically discussed in the text are highlighted in bold).

	Experiment 1		Experiment 2	
	F	***p***	***F***	***p***
Format	**1.96**	**.161**	**22.63**	**< .001**
Nutrient	1.82	.162	0.07	.793
Quantity	3.26	.011	2.90	.089
Correct decision	15.81	< .001	43.28	< .001
Format × nutrient	1.22	.295	1.36	.245
Format × quantity	2.84	.023	0.04	.837
Format × correct decision	**25.72**	**< .001**	**6.32**	**.012**
Nutrient × correct decision	18.57	< .001	61.74	< .001
Quantity × correct response	19.52	< .001	2.21	.138
Nutrient × quantity	1.00	.436	0.09	.771
Format × nutrient × quantity	0.27	.976	0.03	.871
Format × nutrient × correct response	**3.20**	**.041**	**0.37**	**.544**
Format × quantity × correct response	1.84	.119	0.49	.486
Correct response × nutrient × quantity	0.26	.979	6.69	.010

Note. Levels for each fixed effect were as follows: format = 2 (verbal or numerical); nutrient = 3 in Experiment 1 (fat, sugar, or minerals), 2 in Experiment 2 (fat or minerals); quantity = 5 in Experiment 1 (very low, low, medium, high, very high), 2 in Experiment 2 (low or high); correct decision = 2 (healthy or unhealthy).

verbal than numerical quantifiers, as quantified by an interaction between format and correct decision, $F(1, 2733) = 25.72$, $p < .001$. Pairwise comparisons between the correct decision conditions showed that specifically, with verbal quantifiers, participants were more likely to believe an unhealthy combination of minerals was healthy than vice versa, suggesting that participants did not rely on the type of nutrient to make their decision, $F(1, 2733) = 85.53$, $p < .001$, $M_{diff} = 39.52\%$, 95% CI [31.11, 47.90]. They did not make the same error pattern in the numerical condition, suggesting that the context did not factor heavily into their decision, $F(1, 2733) = 2.54$, $p = .111$, $M_{diff} = 6.54\%$, 95% CI [-1.57, 15.25].

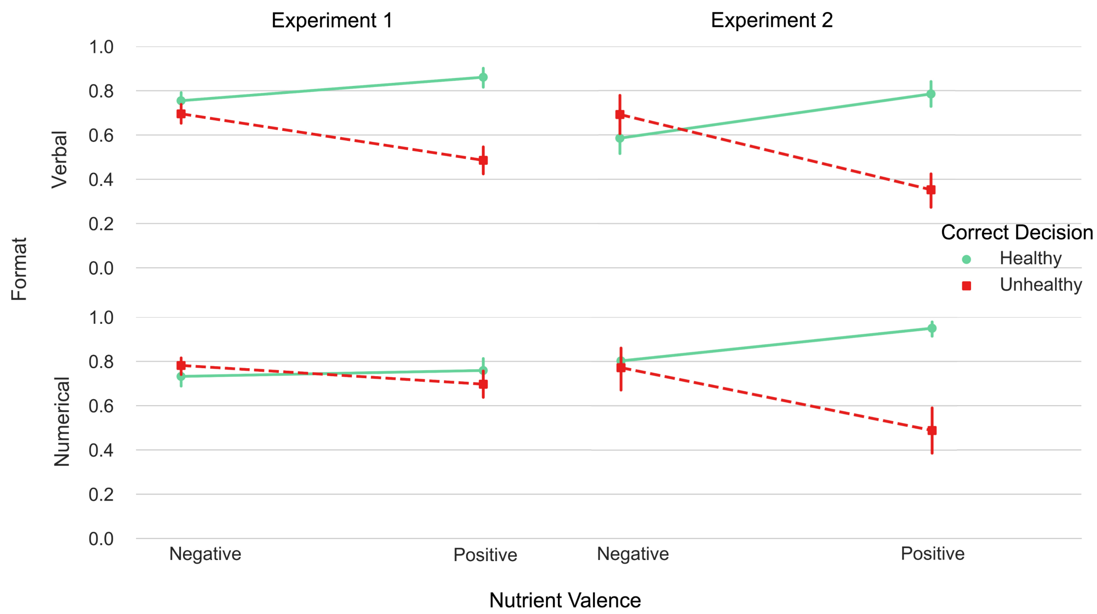

Figure 3.3. Effects of format, nutrient, and correct decision on decision performance in Experiment 1 and 2.

3.3.3 Discussion

Our exploratory analysis into participants' decision patterns suggested that the context provided by the type of nutrient in the verbal labels influenced participants' decisions more than in the numerical labels. This information was not strictly needed to perform the task, since the decision relied on whether the sum of quantities was more or less than 100%, irrespective of the nutrient. The nutrient type simply presents a shortcut to make the decision before performing the full calculation. Participants' pattern of responses, indicating a greater reliance on the nutrient with verbal quantifiers, could therefore be taken as evidence that they used these shortcuts more with verbal than numerical quantifiers. However, the results also showed that on average, response times, performance, and subjective effort did not differ significantly between formats, which could indicate that there was no difference in the processing of the two quantifier formats. Building on these contrasting findings, we planned a second experiment to retest the effect of format on two correlates of processing style: response time, decision performance, and to test the reliance on the nutrient in a confirmatory analysis (rather than with exploratory ones as in Experiment 1).

The method of Experiment 2 was also improved to control for possible variation in the numerical meaning of the verbal quantifiers across participants. One could interpret the findings of Experiment 1 as resulting from a skewed meaning of verbal quantifiers (i.e., participants systematically interpreting verbal labels as less than the matched numerical values). We therefore adapted the method to rule out this possibility and ensure that the tendency to mistake counter-intuitive verbal combinations was not due to interpretational variation.

3.4 Experiment 2

In Experiment 2, we sought to replicate the effects of Experiment 1 for three indicators of processing style: response time, decision performance, and use of context-based shortcuts. We tested the robustness of our findings while addressing two main limitations. First, we provided a planned test of the interaction effects found in Experiment 1, addressing the issue of potentially inflated Type I error rates when relying on exploratory analyses (Wagenmakers et al., 2012). Second, we eliminated the effects of interpersonal variability in numerical interpretations of verbal quantifiers by matching verbal quantifiers to a personalised numerical interpretation for each participant.

3.4.1 Method

Participants. The experiment was powered to detect the format $\times$ nutrient $\times$ correct decision interaction with effect size f = .10 (α= .05, 1-β= .80, two-tailed test based on a mixed ANOVA) after accounting for outliers who might translate verbal quantities into excessively high or low amounts. We obtained data from 154 participants after excluding 11 submissions that failed a pre-registered check for reading attention. The sample was 64% female, 89% White, with ages ranging from 19-71 (M = 36.80, SD = 11.34). Seventy-seven percent had a university degree. Participants had a slightly higher preference for deliberation than intuition (M_{diff} = 0.32, SD_{diff} = 0.71). They had slightly positive attitudes towards healthy eating (M = 4.98, SD = 1.06). Forty-seven percent had a healthy mean estimated BMI (self-reported) and 68% reported general use of nutrition labels.

Design. The design was similar to the one of Experiment 1, with format manipulated between-subjects and the other factors (nutrient, quantity, and correct decision) manipulated within-subjects. We reduced the number of nutrients and quantities to two each (nutrient: fat [negative] and minerals [positive], quantity: low and high), such that each participant completed eight trials in total, with the order of presentation of within-subjects factors randomised.

Materials and procedure. The experiment was programmed using Inquisit 5 (Millisecond Software, 2016) and delivered online through a survey panel (to take part, participants temporarily installed the Inquisit web plugin to their computer). A checking question (whether participants agreed with the statement, '*I am using a computer at the moment.*') was included at the end of the experiment to ensure participants were not responding carelessly.

Participants performed first a verbal-to-numerical translation task, followed by the GDA decision task from Experiment 1.

Verbal-to-numerical translation task. Participants provided numerical percentages for four verbal labels (presented in random order): low % fat, low % minerals, high % fat, and high % minerals. Participants' translations were used as the numerical quantifiers in the decision-making task. Table 3.5 shows the distribution of participants' translations, which were on average lower than in Experiment 1.

Decision-making task. Participants then completed the same GDA decision task as in Experiment 1. The only difference was that the numerical condition of the task used the unique numerical value for low fat, low minerals, high fat, and high minerals that participants had themselves provided in the translation task. In the verbal condition, the quantifier was low or high. Participants read the same instructions about how to decide whether a quantity was healthy or unhealthy, followed by three practice trials, and then the decision trials.

In Experiment 1, we were able to manipulate whether the correct decision

Table 3.5. Descriptive summaries for translations of verbal to numerical quantifiers of fat and mineral labels in Experiment 2.

	Fat		**Minerals**	
	Low	**High**	**Low**	**High**
Mean (*SD*)	10.51% (9.79)	50.29% (27.04)	11.53% (9.71)	56.30% (27.59)
Median	9.00%	50.00%	10.00%	60.00%
Inter-quartile range	10.00%	43.75%	10.00%	50.00%

Note. The large variability in translations was consistent with the literature on interpretations of probability and frequency quantifiers (e.g., Budescu & Wallsten, 1985; Collins & Hahn, 2018; Juanchich et al., 2019).

for the quantity combinations was 'healthy' or 'unhealthy' because the numerical values were set in advance. However, since participants provided the numerical values for Experiment 2, it was less straightforward to create some trials in which the correct decision would be healthy and some in which it would be unhealthy. In order to capture a range of possible correct decision combinations, we used four pie charts derived from the low and high combinations in Experiment 1. These depicted different levels of previous consumption for both low and high quantities (see Table 3.6). The correct decision was healthy if the sum of the pie chart value and the numerical value given by the participant fell within 100%; it was unhealthy if it exceeded 100%. For example, if the participant translated 'high %' as 60%, combined with a previous quantity pie chart of 22.03%, this would be healthy. If the translation were 80%, this would be exceeding limits. Overall, 65% of trials fit within limits and their correct decision was healthy.

We measured decision performance and response time for each of the 8 decision trials. Following our pre-registered protocol, we excluded 76 trials (5.6%) where the numerical translation for a low quantity was equal to or exceeded the translation for a high quantity of that nutrient. We also excluded 15 trials (1.1%) where the response time was below the threshold for manual response to a visually-perceived stimulus (< 150ms; Amano et al., 2006).

Table 3.6. Correct decision condition based on participants' translation of verbal labels and the previously consumed amount.

Pie chart value	**Translation provided by participant**	**Quantity**	**Correct decision**
74.21%	0-25.79%	Low %	Within limits (healthy)
91.13%	0-8.87%	Low %	Within limits (healthy)
22.03%	0-79.97%	High %	Within limits (healthy)
41.65%	0-58.35%	High %	Within limits (healthy)
74.21%	25.79-100%	Low %	Exceeds limits (unhealthy)
91.13%	8.87-100%	Low %	Exceeds limits (unhealthy)
22.03%	79.97-100%	High %	Exceeds limits (unhealthy)
41.65%	58.35-100%	High %	Exceeds limits (unhealthy)

3.4.2 Results

We hypothesised that, in line with more intuitive processing, verbal quantifiers would result in quicker response times, poorer decision performance, and greater reliance on contextual information. As pre-registered, we conducted multilevel analyses using random by-participant intercepts (the full effects model did not converge). As shown in Table 3.2, the speed of participants' decisions was not affected by format, $F(1, 1130) = 1.77$, $p = .184$, $M_{diff} = 0.01$, 95% CI [-0.06, 0.07]. However, participants' decision performance was lower with verbal than numerical quantities, $F(1, 1130) = 22.63$, $p < .001$, $M_{diff} = 18.7\%$, 95% CI [11.53, 25.93] (see Table 3.4)[1].

Participants used nutrient-based shortcuts in their decisions As shown in Table 3.3, the proportion of errors made when the normatively correct response conflicted with an intuitive response for a nutrient was higher for verbal than numerical quantifiers. This was quantified by a significant interaction between format and correct decision, $F(1,\ 1130) = 61.74$, $p < .001$. Although the three-way interaction between format, nutrient, and GDA fit was not significant, $F(1,\ 1130) = 0.37$, $p = .544$, planned pairwise comparisons showed that participants were specifically more likely to judge a verbal quantity of fat as unhealthy than healthy, $F(1,\ 1130) = 3.95$, $p = .047$, $M_{diff} = 11.9\%$, 95% CI [0.15, 23.74]. However, this was not the case when the quantifier was numerical, $F(1,\ 1130) = 57.26$, $p < .001$, $M_{diff} = 2.7\%$, 95% CI [-8.96, 14.31].

3.4.3 Discussion

The results of Experiment 2 supported our prediction that participants would perform worse with verbal than numerical quantifiers, but response times did not differ across formats. We found that participants committed more intuitive errors with verbal than numerical quantifiers. These errors were indicative of their reliance on the nutrient even though it was irrelevant to the decision task. In this experiment, we judged participants' decisions using accuracy criteria that accounted for individual differences in verbal quantifier interpretations, instead of assuming participants to interpret quantifiers in line with psychologically average values. Therefore, the effects on decision performance and use of context-based shortcuts are more likely attributable to the difference in quantifier format rather than to differences in interpretations.

[1]We checked if scoring the verbal quantifier decisions based on criteria for Experiment 1 (i.e., based on the average translations provided in that experiment) would have resulted in better verbal quantifier performance. In fact, using the criteria reduced overall performance from 62% to 38%.

3.5 General Discussion

We aimed to test the effect of quantifier format (verbal or numerical) on processing style in two experiments, using four correlates of processing style derived from dual-process theory: subjective effort, response time, decision performance, and reliance on contextual information (Evans & Stanovich, 2013; Kahneman, 2011). We expected effects for these variables to converge to provide robust evidence that verbal quantifiers are processed more intuitively than numerical quantifiers, but the results did not provide clear-cut evidence. Results varied across variables and experiments. Participants did not respond quicker or more effortlessly with verbal than numerical quantifiers. Although participants did not perform significantly differently between formats in Experiment 1, when we took into account what each participant believed the verbal quantifiers to mean (Experiment 2), we found that participants' decision accuracy was better for numerical than verbal quantifiers. Additionally, there was evidence that people relied more on contextual cues as a shortcut with verbal quantifiers, even when they should not need to use these cues: in both experiments, participants consistently made more errors with verbal than numerical quantifiers when the context did not match the correct decision. This result aligns with traditional views of intuitive processing, where an initial incorrect intuitive response may need correcting from the analytical system (Sloman, 1996)

3.5.1 Are decisions better (but slower) with numerical quantifiers?

Based on dual-process theory, we expected that with numerical quantifiers, people would make better decisions that took longer and required more effort as compared to decisions with verbal quantifiers. The effect of format on decision-making was mixed: on average, participants made better decisions with numerical quantifiers, but this was only significant in Experiment 2. Experiment 1's findings match with previous work that did not report significant differences in average performance between verbal and numerical probabilities (González-Vallejo et al., 1994; Olson & Budescu, 1997) —and also did not control for variability in verbal quantifier interpretation. When we controlled for this variability in Experiment

2, ensuring that the numerical and verbal quantifiers presented were equivalently matched, decisions were significantly better with numerical than verbal quantifiers. This suggests that participants do indeed make better decisions when the same quantity is presented numerically than verbally.

The effect of format on reaction time and effort gave a mixed picture on whether verbal quantifiers were processed more intuitively than numerical ones, as there was no evidence that the formats differed on these measures. Many decision-making models predict that quicker decision times should be associated with a rising error rate (Bogacz et al., 2006). Poorer decision quality, especially on mathematical tasks, is further posited to be a trademark of intuitive as opposed to analytical processing (Kahneman, 2011; Rusou et al., 2013). Yet our results paint a conflicting picture between the different indicators of processing styles, with better performance not associated with longer response times in Experiment 2. In general, participants' decision times were fast (less than 2s), which means they processed numerical quantifiers quickly *and* accurately. This is contrary to the traditional view that intuitive processing produces responses that are quick but often need to be corrected by the slower analytical system (Sloman, 1996), however more recent challenges to the concept of intuitive processing posit that intuition can be accurate under the right circumstances (Bago & De Neys, 2019; Plessner & Czenna, 2008); based on this view, participants could be processing numerical quantifiers intuitively, but accurately. For example, Bago & De Neys (2019) found that people who got logical reasoning tasks correct often already had a correct instinctive response to the question. Other processing models, such as the parallel constraint satisfaction model, suggest that people can process even complex information in surprisingly short times (Glöckner & Betsch, 2008b; Trippas et al., 2017). These network-based models conceptualise a decision process as a series of activations, where over the time-course of the decision activation strengthens for one option over another (Glöckner et al., 2014). In the case of our decision task, a network model might show different activation patterns of verbal and numerical quantifiers, such that even though the time taken to make the decision was similar, the more activated option (healthy or unhealthy differed). By this view, verbal and numerical quantifiers might be similarly intuitive, but the numerical quantifiers might have led to stronger activation of intuitions about quantitative information (more likely to lead to correct responses) while the verbal quantifiers might have led to stronger activation of intuitions about the context (less likely to lead to correct responses).

3.5.2 People used context-based shortcuts more with verbal quantifiers

While decision performance, reaction time, and effort gave a mixed picture of which quantifier format led to more intuitive processing, our other measure, the use of context-based shortcuts, shed more light on the processes underlying participants' decisions. Participants tended to use nutrient information more in decisions with verbal quantifiers, as indicated by a higher proportion of errors in trials where nutrient information conflicted with the normative correct response (e.g., needing to judge a quantity of minerals as unhealthy, or fat as healthy). In the experiments, we specifically designed a decision task (judging whether a combination of two quantities exceeded a limit) that could be completed without having any contextual information about what the nutrient was. If participants performed the task by the given criteria, we would expect similar levels of performance for both types of decisions (the correct decision being healthy or unhealthy). However, we observed more errors in matching minerals to 'unhealthy' (and to a smaller and less consistent extent, fat to 'healthy'), which was consistent with what people would assume about the context. This error pattern suggests a difficulty in suppressing an intuitive answer that is misaligned with the correct decision, and provide evidence for greater intuitive processing with verbal quantifiers (De Neys, 2017a). Participants' patterns of decision-making thus offer support for the proposition that verbal formats elicit intuitive biases (Windschitl & Wells, 1996). In the traditional view of dual-process models, intuitive systems are characterised as relying on shortcuts to make a decision (Kahneman, 2011). The nutrient in our case presented a shortcut to the problem based on existing knowledge of about properties of fat (typically unhealthy) and minerals (typically healthy). This shortcut substitutes for the more onerous process of comparing the quantities to make a decision (Kahneman, 2011; Kahneman & Frederick, 2002).

Participants' decision-making pattern also fit other influential dual-process models, such as the default-interventionist view (e.g., Evans & Stanovich, 2013; Sloman, 1996). This model predicts that when people use intuition, they should make more errors when the intuitive answer conflicts with the normative one (Bago & De Neys, 2017). In our experiments, we presented trials in which participants would be tempted to pair a nutrient with one but not the other answer (e.g, minerals –healthy). On trials where the correct decision conflicted with this automatic association, participants thus had to suppress their existing associations about the nutrient (i.e., the 'intuitive' answer based on the context) in order to answer correctly. This requires a cognitive decoupling of information typically associated with analytical processing (Evans & Stanovich, 2013). We found that in such conflict trials, verbal formats tended to produce a greater proportion of errors (nearly 50% more) than numerical formats. This suggests that verbal formats might recruit intuitive processes more heavily than numerical formats.

3.5.3 Alternative interpretations of results

While a dual-process explanation could explain our findings that verbal quantifiers prompt greater use of context-based shortcuts, it is worth considering whether other theories about differences between verbal and numerical quantifiers could also explain participants' responses, especially since not all of our measures of intuitive processing aligned. Other lines of research posit that verbal quantifiers produce different decisions from numerical ones because they are more vague (Budescu & Wallsten, 1995) and can be interpreted in terms of the base rate of the context (Weber & Hilton, 1990), or convey information about the focal points of the quantifier (Moxey & Sanford, 1986; Teigen & Brun, 2000).

A vagueness explanation would argue that the greater vagueness of verbal quantifiers (Budescu & Wallsten, 1995) resulted in more decision errors because it is difficult to combine imprecise quantities. We believe, however, that this explanation is less likely. First, we depicted previously consumed quantities using a pie chart, which introduced some vagueness in the initially-presented quantity. This would have reduced the level of precision with which participants could make their decision across both formats. Second, vagueness in interpretations of verbal quantifiers should make it more difficult for participants to decide in one type of combination, but also easier to decide in the other. For instance, in Experiment 1, a participant who interpreted 'low' as less than 20% would consistently perform better with a healthy combination, but worse with an unhealthy one, resulting in a similar average performance to a participant who made equal proportions of errors in either direction with the 20% quantifier. Across a sample, we would also expect that participants would make on average the same proportion of each type of error (mistaking healthy or unhealthy combinations). The consistently greater prevalence of errors for cases such as an unhealthy quantity of minerals (as opposed to a healthy quantity of minerals) thus indicates that verbal vagueness is not solely responsible for performance differences.

Another possible explanation is that verbal quantifiers may be interpreted in a relative fashion, as a function of the expected base rate (Weber & Hilton, 1990). In the context of nutrient quantities, this could mean an interpretation of 'high' as being high relative to what one expects. The translation procedure in Experiment 2 was designed to tackle this issue, as participants' translations should already reflect what they believed to be a high % for the given context. We observed similar patterns in performance and error types even after this procedure, which suggests that it is not simply that participants were translating verbal quantifiers as a relative amount. However, it is also possible that base rate expectations about particular nutrients remained salient when verbal quantifiers were presented —for instance, prevalent beliefs that most foods will not contain many minerals, so adding as much as one can to one's diet is good. This is a different type of intuition from a rapid association of minerals with healthiness, but also reflects an additional difficulty in dissociating knowledge about minerals, and could explain why people had more trouble making counterintuitive decisions about minerals than about fat. Because the task was structured such going over the limit was unhealthy, it was better suited to nutrients like fat and sugar, for which people regularly worry about exceeding limits. Future research might consider reversing the criteria such that the 100% is a target to be met, rather than avoided. This could ascertain if the counterintuitive decision for fat would become more difficult relative to minerals, and also remain harder to suppress for verbal than numerical quantifiers.

Finally, one could argue that verbal quantifiers provide a stronger focus on either the nutrient described, or away from it (Moxey & Sanford, 1986; Teigen & Brun, 2000). This focusing property has been shown to be clear for verbal quantifiers (e.g., 'it is likely' leads people to think about what will happen, whereas 'it is unlikely' leads people to think about what will not happen) than for numerical quantifiers (the equivalent numerical probabilities are more ambiguous in what focus they evoke; Teigen & Brun, 2000). If the verbal quantifiers in our task put a focus on the nutrient present, this could also explain people's tendency to rely on the nutrient in decision-making. However, this explanation does not exclude the possibility of participants being more intuitive with verbal quantifiers; rather, it explains *why* participants were more intuitive. If a verbal quantifier's focusing properties encouraged people to use the context as a shortcut, they might then be less likely to perform the task in an analytical manner.

3.5.4 Limitations

The current results were derived from two well-powered pre-registered experiments. However, the methodology relied on quantities that were typically round figures. This might reduce the level of effort required to process them (DeStefano & LeFevre, 2004). Further, both our samples were generally well-educated, which could indicate a high level of numeracy (Parsons & Bynner, 1998), meaning that participants would have found it easier to perform numerical tasks (although education does not always predict numerical ability; Lipkus et al., 2001). Hence, before drawing a firm conclusion from our results and assuming, for example that numerical quantifiers are intuitively processed based on the quick reaction times, future work should test a wider range of numerical values while controlling for individual differences in numeracy.

Another limitation of our research is that we focused on a specific task within a nutrition communication context. Our results showed that people do use salient but less relevant information to inform quantitative decisions, but a further extension of this work would be to test whether this holds across alternative scenarios, for instance in traditional gambling tasks. One could also assess whether the greater susceptibility of verbal quantifiers to context-based shortcuts varies depending on the nature of the contextual information (for example, how strongly positive or negative it is). Given the practical implications for applied communications in health and risk, where there is much debate about using verbal or numerical formats to express quantities (e.g., Berry et al., 2003; MacLeod & Pietravalle, 2017; Peters et al., 2009), this is an important direction for research.

3.5.5 Conclusion

Two experiments showed that participants did not differ on response time and subjective effort when making decisions with verbal or numerical quantifiers. However, decisions based on numerical quantifiers were generally better than those with verbal quantifiers, and people tended to rely more on contextual cues with verbal than numerical quantifiers, even when they did not need those cues to perform the task. Taken together, the evidence suggests that the distinction between processing of verbal and numerical quantifiers is not as clear as previous research posited (Windschitl & Wells, 1996). The reasoning that communicating quantities in numerical format increases effort (Malam et al., 2009; Peters et al., 2009) may need to be revisited. Conversely, one could potentially improve decisions with verbal quantifiers by ensuring contextual cues match the correct decision. .

Chapter 4: The Intuitive Use of Contextual Information in Decisions made with Verbal and Numerical Quantifiers

4.1 Abstract

Verbal and numerical formats (e.g., verbal: 'low fat', or numerical: '20% fat') are used interchangeably to communicate nutritional information. However, prior research implies that verbal quantifiers are processed more intuitively than numerical ones, and are more influenced by contextual information. We tested this hypothesis in two pre-registered experiments measuring four indicators of processing style: (i) response time, (ii) decision performance, (iii) reliance on irrelevant contextual information, which we inferred from participants' decision patterns, and (iv) the level of interference from a concurrent memory task. Participants imagined they had consumed a given amount of a nutrient (represented in a pie chart) and decided whether a new quantity (either verbal or numerical) could be eaten within their guideline daily amount (GDA). The experiment used a mixed design varying format (verbal or numerical), concurrent memory load (no load, easy, and hard load in Experiment 1; no load and hard load in Experiment 2), nutrient (fat and minerals), quantity (low, medium, and high in Experiment 1; low and high in Experiment 2), and the assigned correct response for a trial (within and exceeding limits). Participants were faster and made fewer correct decisions with verbal quantifiers, and they relied more on contextual information (i.e., the identity of the nutrient involved). However, memory load did not impair decisions with verbal or numerical quantifiers. Altogether, these results suggest that verbal quantifiers are processed intuitively, slightly more so than numerical quantifiers, but that numerical quantifiers do not require much analytical processing to reach simple decisions.

4.2 Introduction

Decisions are often made in a complex environment with an abundance of options, differentiated by information presented in differing formats. For example, information about food can be presented using numerical values (e.g., '20%') or as a verbal quantifier (e.g., 'low'). Ideally, the best format to present such quantified information should facilitate informed decision-making while not overtaxing cognitive resources. To use the food choice context as an example, people should be able to accurately perceive nutrient quantities communicated while shopping in an environment with information overload. Unfortunately, there is conflicting evidence on whether existing information formats (e.g., labels indicating the percentage of one's 'Guideline Daily Amount'; hereafter 'GDA', that a food provides) achieves these goals (Campos et al., 2011; Grunert et al., 2010b; Levy et al., 2000; Scammon, 1977). While numerical formats are more precise estimates, numbers on food labels are often difficult to interpret (Campos et al., 2011; Liu & Juanchich, 2018). On the other hand, verbal formats may be intuitively easier to understand (Wallsten et al., 1993), but more vague in meaning (Budescu & Wallsten, 1995) and less carefully considered (Just & Wansink, 2014). There is also evidence that the format of a quantity can lead people to rely on different aspects of the overall information to make their decision (González-Vallejo et al., 1994). This paper presents two experiments that test whether verbal quantifiers are more intuitive than numerical quantifiers, and whether they lead to different decision patterns.

4.2.1 Levels of information processing: Intuitive vs. analytical

When people process information, their thinking can range from intuitive (a more automatic, quick process that often involves mental shortcuts to simplify information) to analytical (a more complex process that operates consciously, slower, and requires more effort; Evans, 2008; Kahneman, 2011). These styles of processing, typically described as 'System 1' and 'System 2' (for an overview of dual-processing theories, see Evans, 2008, or De Neys, 2017b), are posited to explain differences in the processing of verbal and numerical quantifiers: verbal and numerical formats appear to prompt intuitive and analytical processing respectively (Windschitl & Wells, 1996).

Several properties of words and numbers support the proposition that verbal quantifiers could be more intuitively processed than numerical ones (Ayal et al., 2015; Budescu & Wallsten, 1990; Dunwoody et al., 2000; Nordgren et al., 2011; Windschitl & Wells, 1996). In general, words are processed in an automatic manner, needing conscious effort to suppress the meanings they evoke (MacLeod, 1991). In contrast, numbers tend to be processed in a more intentional, algorithmic manner (Tzelgov et al., 1992), which requires more effort (Lan, 2003; Peters et al., 2009). This is not to say that verbal processing is *always* intuitive and numerical processing *always* analytical; indeed, verbal information can be crafted in a complex manner that requires much effort to comprehend (e.g., in verbal reasoning tasks; Evans, 2002), whereas basic comparisons of two numbers in terms of their surface magnitude can be done quickly and intuitively (Viswanathan & Narayanan, 1994). However, in the case of quantified information, people can more easily understand that a verbal quantifier such as 'low' means the amount depicted is small, whereas this is not readily understood from a numerical quantifier such as '20%' (Viswanathan & Childers, 1996).

Other evidence suggests that people might be more susceptible to intuitive biases when processing verbal quantifiers (Welkenhuysen et al., 2001; Windschitl & Wells, 1996). This could lead to poorer decision-making with verbal quantifiers. One might expect incorrect decisions to be naturally due to the vagueness of verbal quantifiers, which tap into a wide range of possible numerical meanings (Budescu & Wallsten, 1985). This could lead to over- or underestimation of an actual quantity that affects decision-making. For example, someone who estimates a high % of fibre to mean 60% might incorrectly assume they have eaten enough fibre if high only means 30% (see Chapter 2). However, this sort of estimation error should have a facilitative effect in cases where, for instance, someone who underestimates the intended meaning of high % minerals would more easily identify correctly when they have eaten too little. As such, assuming people over- and underestimate verbal quantifiers normally around the mean interpretation, vagueness itself should not affect decision-making at the group level. Indeed, some studies have found that people perform similarly at the aggregate level for decisions with numerical and verbal quantifiers (Budescu & Wallsten, 1990; González-Vallejo et al., 1994).

According to dual-process theory, people making decisions based on verbal quantifiers would be expected to make more errors because they rely on their intuition. The type of errors that people make is therefore informative. Intuitive processes lead to reliance on effort-saving decision strategies, such as relying on contextual cues as a substitute to answer a question (Kahneman & Frederick, 2002). For example, people are more influenced by affective information when relying on intuition (Levin & Gaeth, 1988; Slovic et al., 2007). Closer examination of decision performance in past work showed that people given verbal quantifiers were influenced by how positive an outcome would be, as opposed to basing their decision on the value of the quantity when it was presented numerically (González-Vallejo et al., 1994). This suggests that people rely more on the contextual information when they make more intuitive decisions with verbal quantifiers compared to more analytical ones with numerical quantifiers, which could lead to incorrect decisions if the context is not relevant to the decision.

4.2.2 Measuring intuitive and analytical processes using multiple indicators

Identifying intuitive and analytical processing styles is not a straightforward process. Traditional dual-process theories imply that the two processes differ in terms of speed and effort, and the outcome of the processes differ in accuracy (Evans, 2008; Morewedge & Kahneman, 2010). Although the assumption that there are two qualitatively different processes has increasingly been challenged, the core postulates of the theory (that intuitive processing leads to quicker, easier, but less accurate decisions than analytical processing) continue to fuel academic research and influence advice to decision-makers globally (Melnikoff & Bargh, 2018). Direct comparisons between quantifiers and their average numerical translations for measures such as reaction time and decision quality —often measured as indicators of processing style (Evans, 2008; Horstmann et al., 2010) —show that on average, both may be processed in a similar time (see Chapter 3) and lead to similar performance (Budescu & Wallsten, 1990; González-Vallejo et al., 1994). Because response time and performance are contingent on a wide range of factors, the extent to which they reflect processing style is debated. Some dual-process theorists have, for example, argued that analytical processes could be fast (Glöckner & Betsch, 2008b) and intuitive processes could be accurate (Bago & De Neys, 2019). A more stringent manipulation may therefore be necessary to identify the level of processing prompted by verbal and numerical information.

The defining feature of intuition should be its automaticity, in that it does not load working memory (Evans & Stanovich, 2013; but see also Melnikoff & Bargh, 2018, for limitations of this argument). Analytical processing, in contrast, draws on cognitive resources: a person whose cognitive system is loaded with an extra task would have less capacity to process information analytically, and would rely more on intuition in their decision-making (Shiv & Fedorikhin, 1999). Researchers have successfully demonstrated that concurrent cognitive loads impair analytical reasoning, but not intuitive responses (De Neys, 2006).

Building on the assumptions of the dual-process theory and the hypothesis that verbal quantifiers are processed more intuitively and numerical quantifiers more analytically, we expected from Chapter 3 that verbal quantifiers would be processed quicker than numerical quantifiers, and that people would use strategies that rely on contextual information peripheral to the quantitative decision when making decisions with verbal quantifiers (for example, favouring gambles that present larger payoffs, regardless of their probability to win; González-Vallejo et al., 1994). This is in contrast to strategies that rely more on the quantity itself, which we expected when people make decisions with numerical quantifiers. Finally, we expected that manipulating a person's cognitive load should interfere with performance on a decision task based on numerical, but not verbal quantifiers.

4.2.3 Research objectives

The two experiments reported aimed to test the hypothesis that verbal quantifiers are processed more intuitively than numerical ones. To that end, we used a decision task where participants had to judge if a combination of nutrition quantities (presented as 'Guideline Daily Amounts'; or 'GDAs') was within or exceeding a specified limit. This allowed us to set two types of trials: trials where quantities fell within the GDA limit or exceeded it. Thus, participants could make two types of correct decisions (they could be correct that the quantities were within or exceeded the limit) and two types of incorrect decisions (they could be incorrect that the quantities were within or exceeded the limit). We also included different combinations of nutrient and quantity values in the task to create different associative contexts that should suggest different intuitive responses. For example, as illustrated in Figure 4.1, an intuition that the nutrient 'minerals' are healthy (Oakes, 2005a) presents a conflict in a situation where the correct decision is that the quantity exceeds a healthy limit. We measured four indicators of processing style: response time, performance, level of reliance on contextual information, and the effect of interference from a concurrent task. Although response times and performance measures in themselves may not be conclusive evidence for intuitive or analytical processing (Evans & Stanovich, 2013; Horstmann et al., 2010), we also employed a memory load manipulation to tax cognitive resources, which should interfere with performance for analytical, but not intuitive decisions (De Neys, 2006; Trémolière et al., 2014).

Based on our overall hypothesis, we expected quicker and fewer correct decisions with verbal quantifiers, which should also be more influenced by information about the nutrients (context) than decisions with numerical quantifiers. In addition, for the novel test using memory load, we expected that the concurrent cognitive load would decrease performance if a task were analytical. If, in the task, summing the quantities (verbal or numerical) to reach a decision required analysis, memory load should impair correct responding. If it did not require analysis, the memory load manipulation would not have an effect. If, as we expected, the verbal quantifier required less analysis than the numerical, we would see an impairment of the numerical decisions compared to the verbal ones under memory load.

We pre-registered the experimental design, hypotheses, and analyses prior to each experiment. These, along with the materials and data, are available on the Open Science Framework (OSF).

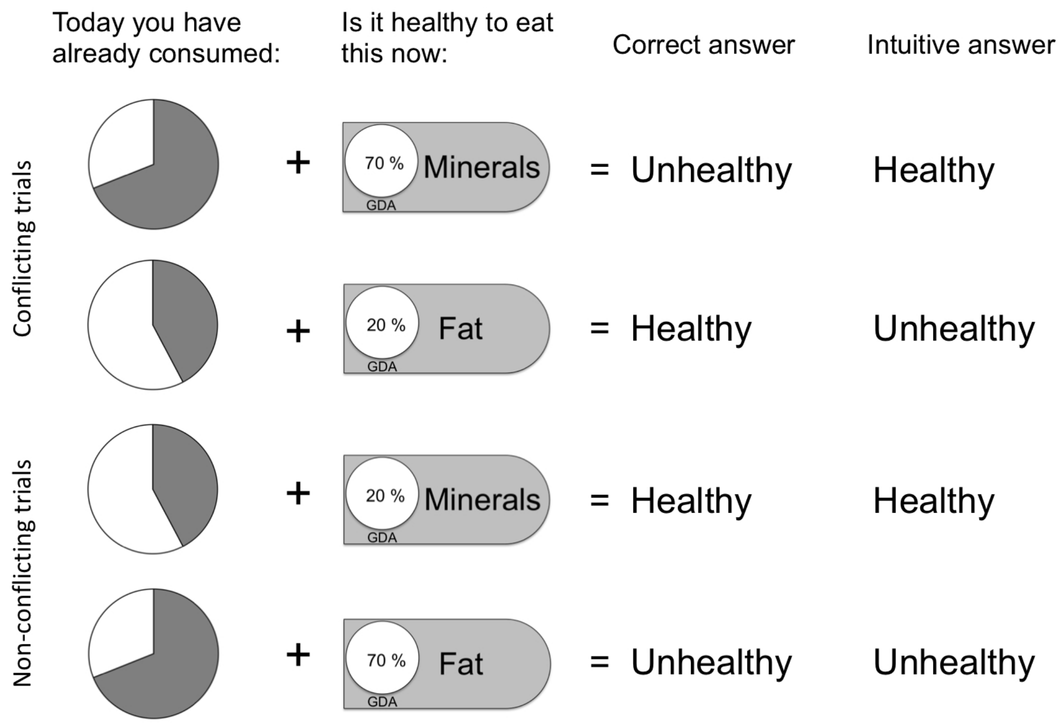

Figure 4.1. Examples of trials where the nutrient could present an intuitive conflict vs. no conflict in the decision task.

4.3 Experiment 1

4.3.1 Method

Participants. Sixty-six participants from a university lab database completed the study (68% female; age range 19-66 years, M = 23.88, SD = 7.90; 52% White, 26% Asian, 17% African; 53% with a university degree). We powered the study to capture a small-to-medium effect for the hypothesised interactions using a mixed variance analysis (Cohen's f = .18, α= .05, 1-β= .80). A sensitivity analysis showed that the recruited sample size had 80% power to detect a medium between-subjects effect of format (f = .25). Participants were paid a £4 show-up fee and given the opportunity to earn additional payment to encourage diligent responding (they were offered £0.10 per correct response on the memory tasks and £0.05 per correct response on the decision tasks).

We measured participants' preferences for intuition and deliberation (Betsch, 2004), their attitudes towards healthy eating (Steptoe et al., 1995), their use of food labels, and Body Mass Index (BMI). Our sample had a preference for deliberation (M = 3.95, SD = 0.50) over intuition (M = 3.42, SD = 0.51), positive eating attitudes (M = 5.15, SD = 1.20) and half reported using nutrition labels regularly. Mean estimated BMI was in the healthy range (M = 22.60, SD = 4.33).

Design. Participants made decisions about whether a given quantity of a nutrient (representing a proportion of their GDA) was healthy to consume given what they had already consumed. We used a 2 (format: verbal or numerical) × 3 (memory load: none, easy, or hard) × 2 (nutrient: minerals or fat) × 3 (quantity: low, medium, or high) × 2 (correct response: within limits —healthy or exceeding limits —unhealthy) mixed design. Format was manipulated between-subjects (random allocation for each participant), while the other factors were manipulated within-subjects (randomly presented across trials). The different combinations of nutrients, quantities, and the assigned correct response allowed us to ascertain the decision strategy participants might use. From a normative perspective, assuming the verbal and numerical quantifiers were strictly equivalent, only information about the quantities should determine if participants decide if it was within limits (healthy) or exceeding limits (unhealthy). The nutrient was not relevant to the decision. However, it allowed us to identify trials that required participants to make a decision that would conflict with an intuitive response to the trial (see Figure 4.1).

Materials. The experiment was delivered using Inquisit 4 (Millisecond Software, 2015; code available on the OSF). There were two task components: the GDA decision task and the memory task.

GDA decision task. To measure decision-making performance, we used a GDA decision task (see Chapter 3). As shown in the top panel of Figure 4.2, in each decision trial, a fixation cross appeared for 500ms, followed by a pie chart illustrating an amount of a given nutrient that participants should imagine they had previously consumed, which was presented for 3000ms. Participants were then presented with a new quantity (either verbal or numerical) of the same nutrient. Their goal was to decide if eating this quantity would fall within their

GDA limit (‘healthy’) or exceed it (‘unhealthy’). They pressed the left arrow key for healthy and the right for unhealthy, or vice versa.

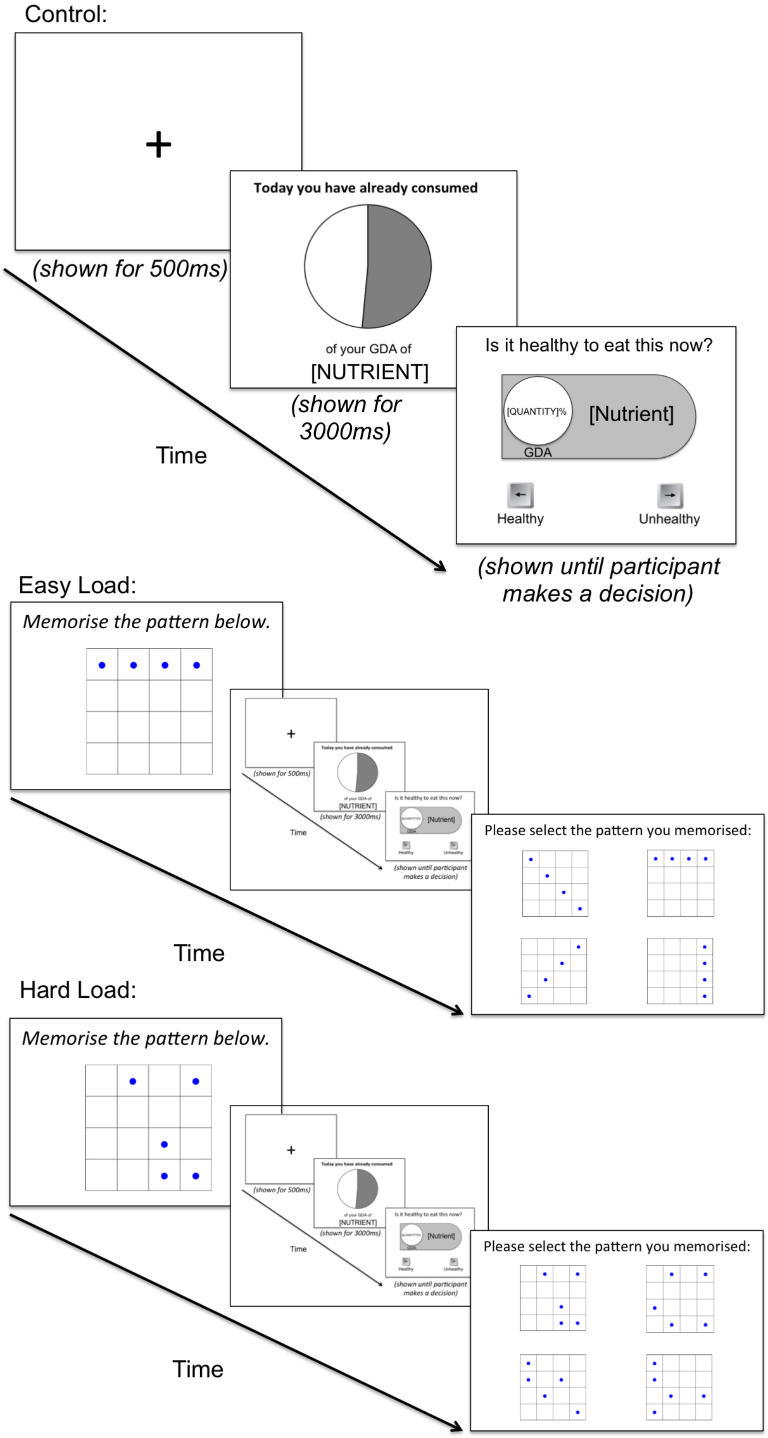

Figure 4.2. Example of a decision-making trial in the no-load, easy load, and hard load conditions in Experiment 1.

Note. The % quantity was either verbal (low, medium, or high) or numerical (20, 40, or 70), and the nutrient was either fat or minerals.

As summarised in Table 4.1, the quantity that followed the initial nutrient intake was either low, medium, or high. Following similar procedures in developing comparable verbal and numerical conditions between quantity formats (Teigen & Brun, 2000; Welkenhuysen et al., 2001), we used corresponding quantities for the two conditions that had been found to be on average psychologically equivalent with similar samples in a similar context (see Chapter 2). The correct response in the task was determined by whether the two quantities added together fell within or exceeded 100% (of the GDA for this nutrient). The amount already consumed (shown in the pie chart) was set such that half the combinations were within the limit and half exceeded it. Based on this design, participants could make two types of correct decisions (they could be correct that the quantities were within or exceeding the limit) and two types of incorrect decisions (they could be incorrect that the quantities were within or exceeding the limit).

Table 4.1. Quantity combinations used in the GDA decision task in Experiment 1.

Amount already consumed	**Decide if eating this quantity is within the GDA limit:**		**Correct response**
	Verbal	**Numerical**	
66.98%	Low %	20%	Within limit (healthy)
44.79%	Medium %	40%	Within limit (healthy)
12.22%	High%	70%	Within limit (healthy)
91.13%	Low %	20%	Exceeds limit (unhealthy)
68.95%	Medium %	40%	Exceeds limit (unhealthy)
51.46%	High %	70%	Exceeds limit (unhealthy)

Memory load manipulation. To manipulate memory load, we used a dot memorisation task (Białek & De Neys, 2017; Trémolière et al., 2014). Participants memorised a dot pattern in a 4×4 matrix (see the middle and bottom panels of Figure 4.2) presented for 2s before performing the GDA decision task. After they made their GDA decision, they selected which of four matrices had been presented. They were told whether their selection was correct. If they erred, they were instructed to try harder on the next trial. There were two memory load conditions, taken from (Białek & De Neys, 2017). In the easy load, four dots were arranged in a straight line, whereas in the hard load, five dots were interspersed. Of the three incorrect matrices, one was more highly similar to the correct one than the others (e.g., sharing three out of five dots). Previous work has established that this is a demanding secondary task that interferes with analytical but not intuitive processes (Białek & De Neys, 2017; De Neys & Schaeken, 2007; Trémolière et al., 2014). The simple pattern minimally burdens cognitive resources whereas the hard one further interferes with analytical reasoning (Białek & De Neys, 2017). Further, we expected the visuo-spatial nature of the load to have a similar impact on analytical processing of either quantifier format.

Procedure. After providing informed consent, participants read the generic rules of the decision and memory tasks. Participants first practised the decision task and had to perform the final of the three practice trials correctly to move on. To reduce learning effects, they received feedback in these practice trials, but not in the experimental trials. Next, participants practised three trials of the memory load task with a blank screen of 500ms between memorisation and recognition. They had to perform the final practice trial correctly to proceed, otherwise they received more practice trials. Before the experimental phase began, they were informed that they could earn £0.05 per correct response on the GDA decision task and £0.10 per correct response on the memory task.

Participants were randomly assigned to either the verbal or numerical version of the decision task. Participants performed 3 blocks of 12 trials each, corresponding to the no-load, easy load, and hard load conditions (see Figure 4.2). The order of presentation of these three conditions was randomly assigned. Within each block, participants made decisions for the 12 decision situations resulting from the randomised crossing of the 3 quantities, 2 nutrients, and 2 assigned correct response manipulations. Participants were given a break at the end of each block. When they had completed all three blocks, they provided a numerical percentage for the three verbal quantifiers, and selected which of five verbal quantiifers (very low –very high) best fit the three numerical quantifiers. This was to check if participants' natural interpretations of the two quantifier formats were psychologically equivalent. Finally, participants provided demographic information.

4.3.2 Manipulation checks

Memory load manipulation check. Memory performance was good overall, with participants selecting the correct matrix significantly more for easy grids (91.2%) than for hard grids (87.2%), $F(1, 1582) = 6.28$, $p = .012$. Participants also took longer to select the hard matrices than the easy ones, $F(1, 1582) = 205.57$, $p < .001$. Cases where participants failed to select the correct grid could indicate that they had not sufficiently burdened their cognitive resources while performing the GDA decision task. Therefore, we excluded all trials where participants selected neither the correct grid nor its close target (which indicated a reasonable memory error even when participants were diligently memorising the grid; Białek & De Neys, 2017 [1]).

Numerical interpretation equivalence check. The mean numerical percentages associated with low, medium, and high verbal quantifiers were close to the numerical quantifiers used in the decision task: 17% vs. 20%, 36% vs. 40%, and 58% vs. 70% respectively. The mean verbal-numerical translations varied widely and were generally right-skewed, with *SD*s of 12%, 14%, and 23% for low, medium, and high. The modal translations were 20%, 50%, and 70%. Translations of the numerical quantifiers (20%, 40%, and 70%) to verbal ones were low, medium, and high respectively (except for 70% fat, for which the verbal translation was 'very high'). We followed up with a logistic regression to ascertain

[1] This procedure was not part of our pre-registered protocol and was suggested by a reviewer. Employing it did not substantially change the results of our analysis.

if tendencies to under- or overestimate verbal quantifiers might result in participants selecting 'healthy' or 'unhealthy' more often (which would result in errors due to translation rather than processing style). This analysis found no significant effect of under- or overestimations on all decisions, all p's $> .100$. A full report of the analysis is included in Appendix C.

4.3.3 Results

To test the effect of format on response time, decision performance, contextual information use, and load impairment, we performed a multilevel model at trial level for response time and decision performance. As response times displayed significant positive skew (original skewness = 3.00), these were log-transformed prior to analysis (resulting skewness = 0.43). We ran the pre-registered statistical model including all two- and three-way interactions, and then a simpler model that better targeted the hypothesised interactions, to avoid Type I error rate inflation (Cramer et al., 2016). The two models provided the same evidence regarding our hypotheses. We report here the results of the second one (see Table 4.2). Results of the full model are presented in Appendix C (Table C.4). The model reported here included fixed effects for format, load, nutrient, quantity, assigned correct response, and the interactions for format × load, format × nutrient, format × quantity, format × assigned correct response, format × nutrient × assigned correct response, and format × quantity × assigned correct response. The analyses were performed in SPSS using a variance components matrix. The full random effects model did not converge, thus we removed random slopes until a convergent model was obtained, which included by-participant random intercepts and random slopes for quantity. Follow-up pairwise comparisons for these effects can be found in Appendix C (Tables C.1 to C.3).

Evidence for more intuitive processing of verbal quantifiers. Three of our measures showed more intuitive processing of verbal than numerical quantifiers. In line with our hypotheses, participants made slower decisions and gave more correct responses with numerical than verbal quantifiers (response time in seconds: $M_{\text{numerical}} = 2.53$, $SD = 2.02$, $M_{\text{verbal}} = 2.03$, $SD = 1.75$; percentage of trials correct: $M_{\text{numerical}} = .82$, $SD = .38$, $M_{\text{verbal}} = .71$, $SD = .46$), F(1,

Table 4.2. Effects of format, cognitive load, nutrient, quantity, and assigned correct response on response time and performance (analysed in multilevel models) in Experiments 1 and 2. Effects specific to our hypotheses are marked with ˆ.

	Response time (log)				Performance (% correct)			
	Experiment 1		Experiment 2		Experiment 1		Experiment 2	
	F	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Main effects								
Format (verbal/numerical)ˆ	8.74	.003	0.39	.533	17.19	< .001	72.78	< .001
Loadˆ	1.54	.214	65.20	< .001	0.28	.757	0.64	.422
Nutrient	40.35	.557	7.70	.006	3.61	.058	10.21	.001
Quantity	4.02	.018	4.49	.034	6.14	.002	44.58	< .001
Correct response	34.53	< .001	17.91	< .001	127.39	< .001	206.49	< .001
Interactions								
Format × loadˆ	0.03	.974	0.35	.553	0.72	.487	0.04	.843
Format × nutrient	0.09	.759	-	-	0.94	.333	-	-
Format × quantity	3.52	.030	0.28	.598	3.88	.021	0.54	.463
Format × correct response	1.14	.285	-	-	0.78	.376	-	-
Nutrient × correct response	-	-	33.75	< .001	-	-	207.71	< .001
Quantity × correct response	-	-	1.28	.258	-	-	5.91	.015
Format × nutrient × correct responseˆ	2.84	.059	0.45	.718	14.92	< .001	4.69	.003
Format × quantity × correct responseˆ	19.20	< .001	-	-	17.61	< .001	-	-
Format × load × nutrient × correct responseˆ	-	-	1.96	.068	-	-	0.22	.969

Note. The error *df* was 2,241 for response time and 2,256 for performance in Experiment 1, and 6,281 in Experiment 2. Reported effects are the main effects and hypothesised interactions specified in the pre-registrations. (Cells marked with a '-' are effects that were not mentioned in the pre-registration.) Effects specific to our hypotheses are marked with ˆ.

2241) = 8.74, p = .003 (response time), F(1, 2256) = 17.19, p < .001 (decision performance). We also found evidence that participants relied more on associative processes and hence used irrelevant contextual information to decide in the verbal than the numerical condition. Because each trial had an assigned correct response, we could infer the type of error participants made based on the variables that interacted with the assigned correct response. For instance, a three-way interaction between format, nutrient, and assigned correct response could indicate that participants were mistaking the quantities to be within the GDA limit for one nutrient with verbal but not numerical quantifiers. Because the nutrients were either associated with healthiness (minerals) or unhealthiness (fat; Oakes, 2005b), we could identify if the mistakes matched a decisional conflict with these associations. Indeed, participants had more trouble making conflicting decisions in the verbal format than the numerical one (see Table 4.3), F(1, 2256) = 14.92, p < .001 (interaction with nutrient); F(2, 2256) = 17.61, p < .001 (interaction with quantity). In particular, the interaction with nutrient was a strong indication of how much context influenced decision-making in either format. Pairwise comparisons showed that participants had more trouble judging mineral quantities that exceeded (i.e., 'unhealthy') than mineral quantities that fell within the limit (i.e., 'healthy') when the quantifiers were verbal than numerical, F(1, 2256) = 28.86, p < .001 (unhealthy minerals); F(1, 2256) = 4.16, p = .042 (healthy minerals). This suggested the use of a 'minerals are healthy' strategy that was more evident with verbal quantifiers. However, the converse prediction, that people would use a 'fat is unhealthy' strategy, was not observed. Participants were more likely to judge quantities of fat as unhealthy than healthy, and they did so more accurately with numerical than verbal quantifiers, F(1, 2256) = 8.47, p = .004 (healthy fat); F(1, 2256) = 8.33, p = .004 (unhealthy fat) [2].

[2] We also ran pre-registered secondary Bayesian analyses to quantify the support for the interaction and pairwise comparisons. We implemented a mixed BANOVA in JASP (default priors, r scale = 0.5). The evidence for the model with a three-way interaction vs. one without it was inconclusive, BF_{10} = 0.81. However, Bayesian t-tests found extreme evidence that participants were more likely to err when required to judge minerals as exceeding limits (unhealthy) in the verbal than the numerical condition, BF_{10} = 104.41. There was only anecdotal evidence in

Mixed evidence for analytical processing of numerical quantifiers. Our fourth measure of processing, cognitive load, did not show the expected effect. We predicted the memory load would result in dampened performance in the numerical condition (expected to require analytical processing), as compared to unchanged performance in the verbal condition (expected to be intuitively processed). Such a pattern of results entailed an interaction effect between format and load, which was not statistically significant, $F(2,\ 2256) = 0.72$, $p = .487$. Further, load did not affect overall performance, $F(2,\ 2256) = 0.28$, $p = .757$, suggesting that participants were intuitive for both formats.

4.3.4 Discussion

Experiment 1 investigated four indicators of processing style that provided mixed evidence for a processing difference between verbal and numerical quantifiers. Supporting the hypothesis that verbal quantifiers would be more intuitively processed, participants were quicker, but made fewer correct decisions with verbal than numerical quantifiers. Participants also relied more on associative thinking with verbal and numerical quantifiers, as they used irrelevant cues to guide their decision. Specifically, they were more prone to deciding that verbal (as compared to numerical) mineral quantities were within limits (healthy). However, cognitive load did not impair decision-making more in the numerical than the verbal condition. For both quantifiers, decision performance was not significantly different under memory load, suggesting that both were similarly intuitive. To assess the robustness of our findings, we aimed to replicate Experiment 1, but with a modification. In Experiment 1, we determined equivalent verbal and numerical quantifiers pairs (e.g., low and 20%) based on data from Chapter 2. In Experiment 2, we addressed the possibility of individual variation in translations by piping participants' numerical translations of the verbal quantifiers into the numerical decision task. processed without much analysis.

favour of no differences between formats in performance when asked to judge fats as within limits (healthy), $BF_{10} = 0.78$.

Table 4.3. Decrease in performance (% of correct answers) between trials where the correct decision was intuitive and when it was not.

Correct decision	**Experiment 1**		**Experiment 2**	
	Verbal	**Numerical**	**Verbal**	**Numerical**
Intuitive: Fat = Unhealthy	62.21%	72.83%	69.33%	73.70%
Counter-intuitive: Fat = Healthy	80.46%	90.18%	56.88%	75.43%
Difference in performance (Intuitive - counter-intuitive)	-18.25%	-17.35%	12.46%	-1.73%
Intuitive: Minerals = Healthy	90.97%	94.78%	83.87%	90.79%
Counter-intuitive: Minerals = Unhealthy	48.48%	72.18%	31.24%	53.31%
Difference in performance (Intuitive - counter-intuitive)	42.49%	22.60%	52.63%	37.48%

Note. A negative performance difference indicates that participants performed better for trials that conflicted with the intuitive response.

4.4 Experiment 2

The goal of Experiment 2 was to replicate the findings from Experiment 1 using the same measures of processing style (response time, decision performance, contextual information use, and interference effect of cognitive load), while accounting for individual variability of verbal quantifiers. To this end, we had participants provide their own interpretations of the verbal quantities of fat and minerals, and used these values in the task, as well as to assess the accuracy of their decisions. To streamline the experimental protocol, we also reduced the number of quantity and load conditions to two each. We pre-registered an analysis model that was targeted towards our three pre-registered hypotheses. First, we predicted that people would make faster and worse decisions with verbal than numerical quantifiers. Second, we predicted that participants would rely more on irrelevant contextual cues to make decisions based on verbal quantifiers. Third, based on the assumption that verbal quantifiers would require less analytical processing than numerical quantifiers, we predicted that verbal quantifiers would be less affected by the addition of a concurrent cognitive load as compared to numerical quantifiers. The pre-registration for the experiment is available on the OSF.

4.4.1 Method

Participants. Based on the effects obtained in Experiment 1, we determined a priori that a minimum sample of 285 participants was required to achieve 80% power to detect a between-subjects format effect with $\alpha = .05$. As the correct response for a trial depended on participants' translations of verbal quantifiers in this experiment, we included a provision in case certain participants were outliers in their translations (expected to be no more than a third of the sample). We therefore targeted 426 participants from Prolific Academic. After excluding all participants who did not meet the pre-registered exclusion criteria, we had a sample of 420 participants (56% female; age range 18-74, $M = 37.79$, $SD = 12.82$; 91% White, 57% had at least a university degree). A sensitivity analysis using 1,000 simulations of the multilevel model in R gave 93% power to detect the main between-subjects format effect based on this sample size. Participants were paid £1.25 to take part in the study, with the opportunity to earn bonus payments based on their performance (£0.05 per correct memory task response and £0.03 per correct decision task response).

Design. Participants performed the same decision task as Experiment 1 in a 2 (format: verbal or numerical) $\times$ 2 (memory load: none or hard) $\times$ 2 (nutrient: minerals or fat) × 2 (quantity: low or high) × 2 (previously consumed amount) design. Format was manipulated between-subjects (random allocation for each participant), while the other factors were manipulated within-subjects (random presentation across trials). The two previously consumed amounts per quantity (see Table 4.4) allowed us to determine the correct response for the trial based on each individual participants' translation of the verbal quantifiers.

Material and procedure. The experiment was delivered using the web version of Inquisit (Millisecond Software, 2016; code available on the OSF). We added a translation element to the start of the experiment: after participants provided informed consent and read an explanation about GDAs, they provided their numerical interpretations (as a percentage) for each of these four quantities: low % fat, low % minerals, high % fat, and high % minerals[3].

Subsequently, the procedure and materials were the same as Experiment 1, except that there was no easy load block and no medium quantities, and the numerical decision trials used participants' provided translations.

GDA decision task. We used the same task as Experiment 1, as illustrated in the top panel of Figure 4.2. However, we defined the correct answers to each quantity combination based on participants' provided translations. As shown in Table 4.4, if the sum of the pie chart quantity and participants' verbal-numerical translation exceeded 100%, the correct decision should be that the new quantity exceeded limits, and was thus unhealthy. For example, if a participant translated 'low %' as 10%, combined with a pie chart value of 91.13%, the quantities would exceed the GDA limit ('unhealthy'), and the participants' response would be scored as correct if they decided it was unhealthy. In this example, if the translation were 5%, it would be within limits, thus a correct response would be 'healthy'. Overall, 67% of trials had the correct response as being within limits. This indicated that as anticipated, approximately one-third of the sample gave values that always added up with the prior nutrient consumption (shown in the pie chart) to be within the GDA guidelines and hence considered within limits,

[3]Overall, participants translated verbal quantifiers into lower values than in Experiment 1 (M_{low} = 10.11%, *SD* = 7.43; M_{high} = 56.48%, *SD* = 21.46).

and healthy (sum of the two quantities ≤ 100% of the GDA).

Table 4.4. Quantity combinations for the decision trials in Experiment 2 (eight per participant), as determined by the value of participants' verbal quantifier translations and the amount shown in the pie chart.

Amount already consumed	**Decide if eating this quantity is within the GDA limit:**		**Correct response**
	Verbal	**Numerical quantifier (provided by participant)**	
74.21%	Low %	0-25.79%	Within limit (healthy)
91.13%	Low %	0-8.87%	Within limit (healthy)
74.21%	Low%	25.79-100%	Exceeds limit (unhealthy)
91.13%	Low %	8.87-100%	Exceeds limit (unhealthy)
22.03%	High %	0-77.97%	Within limit (healthy)
41.65%	High %	0-58.35%	Within limit (healthy)
22.03%	High%	77.97-100%	Exceeds limit (unhealthy)
41.65%	High %	58.35-100%	Exceeds limit (unhealthy)

Memory load manipulation. We used the same load manipulation and procedure as Experiment 1, except that we did not include an easy load condition. Participants selected either the correct grid or its close target on 94% of the trials. We dropped the remaining 6% of trials with neither a correct nor close-to-correct answer, because failing to remember the grid indicates that participants did not pay enough attention to the memory task and hence their cognition might not have been sufficiently burdened during the GDA decision task (Białek & De Neys, 2017).

4.4.2 Results

Following our pre-registered protocol, we dropped data from 15 trials ($<$ 1%) where participants made a decision in less than the threshold for manual response to a visual stimulus (150ms; Amano et al., 2006), and two trials for which the response time was more than 5 *SD* above the mean. We performed a multilevel model at trial level for response time (log-transformed due to significant positive skew; original skewness = 23.39, resulting skewness = 0.48) and decision performance.

In order to test our pre-registered hypotheses, we included the following fixed effects in the multilevel model: main effects of format, load, nutrient, quantity, and correct response, and interactions for format × load, format × quantity, nutrient × correct response, quantity × correct response, format × nutrient × correct response, and format × load × nutrient × correct response. We ran the analyses in SPSS, using a variance components matrix. The full random effects model did not converge, hence we dropped random slopes until we identified a convergent model, which included by-participant intercepts and random slopes for quantity. The results of the analyses are reported in Table 4.2.

Evidence for intuitive processing of verbal quantifiers. Participants again made more correct decisions with numerical than verbal quantifiers (percentage of trials correct: $M_{\text{numerical}} = .76$, $SD = .43$; $M_{\text{verbal}} = .62$, $SD = .49$), although we did not find that they did so significantly more slowly (response time in seconds: $M_{\text{numerical}} = 1.89$, $SD = 2.16$; $M_{\text{verbal}} = 1.84$, $SD = 1.91$), $F(1, 6281) = 72.78$, $p < .001$ (performance); $F(1, 6281) = 0.39$, $p = .533$ (response time).

In terms of reliance on contextual information, we were primarily interested in how the nutrient (which contextualised the quantity) would affect decision performance, despite it being irrelevant to the decision. Participants used the valence of the nutrient to guide their decision: they were more likely to incorrectly decide that the 'good' nutrient (minerals) quantity fell within limits (i.e., was healthy) when it did not, and that the 'bad' nutrient (fat) exceeded limits (i.e., was unhealthy) when it did. This effect was supported by a three-way interaction of format $\times$ nutrient $\times$ correct response, showing that participants used this strategy in their decisions more for the verbal than numerical quantifiers, $F(1, 6281) = 4.69$, $p = .003$. Table 4.3 illustrates the greater performance impairment caused by relying on the nutrient in the verbal than numerical condition, $F(1, 6281) = 58.98$, $p < .001$ (minerals); $F(1, 6281) = 55.28$, $p < .001$ (fat).

Mixed evidence for analytical processing of numerical quantifiers. Decision performance was not more impaired by cognitive load in the numerical condition compared to the verbal one, $F(1, 6281) = 0.04$, $p = .843$. Load also did not impair overall performance, suggesting that numerical quantifiers did not draw heavily on analytical cognitive resources, $F(1, 6281) = 0.64$, $p = .422$.

4.4.3 Discussion

Experiment 2 showed that the general pattern of results found in Experiment 1 persisted even when we accounted for individual variation in participants' translations of verbal quantifiers. Although participants were not significantly faster, they performed worse in the decision task with verbal and numerical quantifiers, and their pattern of errors was in line with the prediction that they would be more affected by contextual information (i.e., the identity of the nutrient) with verbal and numerical quantifiers. However, consistent with Experiment 1, we did not find evidence for a difference in performance under memory load between the conditions. Therefore, only three out of four of our hypotheses were supported.

4.5 General Discussion

The study investigated whether verbal quantifiers were processed more intuitively than numerical ones in a decision task that required participants to decide if a combination of two nutrient quantities fell within a healthy limit. As single measures (e.g., response times) often cannot provide conclusive evidence of processing styles (Bago & De Neys, 2017), we used four indicators to identify intuitive processes: faster responses, lower decision performance, greater use of irrelevant contextual information, and a lack of interference from cognitive load, with the latter being the critical test of processing style. We expected participants to display these indicators of intuitive processing for decisions with verbal quantifiers more than numerical quantifiers. However, results were mixed. Verbal quantifiers led to fewer correct decisions and greater reliance on irrelevant contextual cues in both experiments, but verbal quantifiers led to faster decisions only in Experiment 1. Finally, the memory load did not affect decision performance for either verbal or numerical quantifiers.

4.5.1 Are both verbal and numerical quantifiers intuitive?

Evidence for intuitive processing of verbal quantifiers. Across all four measures of processing style, both experiments found evidence that participants completed the verbal decision task intuitively. Participants made their decisions quickly (around 2s) and their accuracy was not much above chance. The data also showed that participants relied on irrelevant contextual information to make their decision, for instance not overriding the conflicting association that 'minerals are healthy' when identifying an exceeded quantity of minerals. More critically, their decisions remained unchanged under memory load, which we expected to tax performance only if analytical processing were required (Evans & Stanovich, 2013).

Mixed evidence for intuitive processing of numerical quantifiers. The evidence for whether numerical quantifiers were analytically or intuitively processed was mixed. Compared to the verbal condition, numerical quantifiers appeared less intuitive on three measures: participants made more correct decisions in the numerical than verbal condition, and they did so slower, although the pattern of slower responses was only significant in Experiment 1. They relied less on the irrelevant context, showing a greater ability to overcome associative conflicts in the decision task. However, it is important to note that this could also be because numerical quantifiers are less prone to such conflicts because they are less easily integrated with context than verbal quantifiers (Sanford et al., 1994). For instance, previous work found that participants remembered more contextual information presented with verbal than numerical quantifiers (Moxey & Sanford, 1993). Our critical test of processing style was the effect of memory load, and this did not differ across verbal and numerical formats. The fact that decision performance remained similar in both loaded and unloaded conditions suggests that participants did not use more analytical effort in the numerical condition.

Our findings support previous suggestions (Windschitl & Wells, 1996) that verbal quantifiers elicit intuitive processes, but not that numerical quantifiers elicit analytical ones. This seems surprising, since research from various domains report that numerical information is effortful to process (e.g., nutrition, Campos et al., 2011; healthcare, Peters et al., 2009; medical risks Edwards et al., 2002). This may, however, depend on the specific numerical quantities used. Numerical processing shows greater impairment under a concurrent load if the arithmetic task is more difficult (DeStefano & LeFevre, 2004). In both our experiments, numerical values tended to be rounded to the nearest ten, even those provided by participants in Experiment 2. These values might have been easier to process arithmetically. It is possible that more complex numerical values (e.g., non-rounded values such as 73% instead of 70%; Jaffe-Katz et al., 1989) would draw further on analytical processes and thus be affected by memory load.

4.5.2 Implications for theories of quantifier processing

We derived our hypotheses from the basic, dichotomous dual-process model as a direct empirical test of processing differences between the formats within this framework, which assumes that intuition is fast, does not load on working memory, and is prone to errors and biases (De Neys, 2017b). Critiques of dual-process theory point out that response times and performance are insufficient on their own as indicators of processing style because intuition is not always inaccurate (Plessner & Czenna, 2008) and correct decision outputs that were traditionally classified as analytical can proceed quickly (Bago & De Neys, 2017; Glöckner & Betsch, 2008a). Our findings corroborate this perspective: in particular, the better decisions participants made with numerical than verbal quantifiers did not align with a consistently slower decision time, nor impedance from the memory load. This suggests that in some contexts, people can produce better answers without compromising decision speed. A more recent dual-process model conceptualises intuition as a process that produces both logical and heuristic responses initially, with analytical processing triggered if one detects a conflict between these responses and decides to investigate further (Pennycook, 2017). Applying this to numerical and verbal quantifiers, we see a possibility that a different intuitive response could be generated for each: a logical response for numerical quantifiers (based on the quantity) and a heuristic one (based on the context) for verbal quantifiers. Further, Bago & De Neys (2019) posit that the role of analytical processing may not be to correct a mistaken intuitive response, but to rationalise and support one's initial answer. Indeed, this sort of post-hoc justification of an initial decision does occur when people make food choices (Rayner et al., 2001). A final decision could therefore reflect a multi-step process in which aspects of the information compete in parallel to influence the decision (Busemeyer & Johnson, 2007). A choice between two foods, for instance, can depend on the accumulation of value signals on a sensory (e.g., taste) and a judgemental (e.g., healthiness) dimension, with healthiness accumulating slower than taste (Sullivan et al., 2015). It is possible that in the GDA decision task, where the objective was to judge a combination of quantities, the verbal format accumulated evidence quicker for the holistic goal (whether consumption was healthy), whereas the numerical format accumulated evidence quicker for the rule-based goal (consumption is healthy only if it does not exceed 100%).

Our two experiments also found a greater use of contextual cues in decision-making with verbal than numerical quantifiers, which further informs the difference in processing between the two quantifier formats. A traditional view of verbal quantifiers is that their vagueness impairs decision performance (Berry et al., 2004;

Huizingh & Vrolijk, 1997; Mazur et al., 1999; Visschers, 2008). Our findings show that it is not just verbal vagueness driving this effect. First, we found that participants were less correct with verbal than numerical quantifiers even when we adjusted the numerical values and accuracy criteria to account for variations in participants' translation of verbal quantifiers. Second, misinterpretation of verbal quantifiers cannot explain why participants would make a certain type of incorrect decision. When the quantifier was verbal (compared to numerical), participants relied more on the nature of the nutrient rather than on the quantity itself to assess whether eating it would exceed their daily limit. For example, a verbal quantity of a desirable nutrient (minerals) was more often judged a within limits when it actually exceeded limits. Thus, intuitions based on the learned associations of the nutrients with healthiness or unhealthiness (Oakes, 2004; Wansink & Chandon, 2006) intruded on a task where the nutrient should not have affected the decision.

4.5.3 Implications for food decision-making

Testing whether verbal quantifiers are indeed processed more intuitively than numerical ones is not only relevant from a theoretical and empirical perspective. At an applied level, it is also consequential because efforts to simplify consumer information (e.g., on nutrition labels) have been premised on verbal labels being less difficult to process than numerical ones (Cowburn & Stockley, 2005). Research has also shown that people often rely on mental shortcuts to make food judgement and choices (Gomez, 2013; Scheibehenne et al., 2007; Schulte-Mecklenbeck et al., 2013). Using shortcuts based on contextual information for verbal more than numerical quantifiers thus has further implications on everyday food decisions: a greater tendency with verbal quantifiers to judge unhealthy amounts of 'good' food as healthily within their consumption limit could lead to overconsumption of these foods (Ebneter et al., 2013; Gravel et al., 2012; Wansink & Chandon, 2006). Our findings suggest that numerical quantifiers are less susceptible to these contextual influences, but contrary to previous beliefs (Malam et al., 2009), do not necessarily require more effort or time to process. Numerical quantifiers might thus still be better at facilitating healthier eating decisions.

4.5.4 Conclusion

Our results indicate that when deciding whether a nutrient quantity was a healthy addition to one's daily diet, verbal quantifiers were processed intuitively: participants made quicker and less correct decisions that relied on irrelevant contextual cues, and their ability to make decisions was not impaired when their working memory capacity was diminished. We predicted that numerical quantifiers would differ and be processed more analytically, but the evidence for this was more mixed. While participants were slower, more correct, and used less irrelevant information in their numerical decision-making, they were not impaired by a memory load. This suggests that contrary to previous assumptions, numerical quantifiers may result in quicker *and* more correct decisions.

Chapter 5: Eye-tracking Evidence for Attention Asymmetries in Verbal and Numerical Quantifiers

5.1 Abstract

When people make decisions involving quantifiers, they are affected by the format in which this information is communicated. For example, prior research shows that an attribute with a numerical quantifier (e.g., '5% fat') is evaluated differently when it has a verbal quantifier (e.g., 'low fat'). In a large, pre-registered eye-tracking experiment ($N = 148$), we investigated whether (i) people paid more attention to numerical than verbal quantifiers on a simple nutrition label; (ii) whether attention to positive and negative attributes was different between verbal and numerical quantifiers, and (iii) people's judgements of the label. Participants' eye movements were tracked as they judged the healthiness of 48 labels with a nutrient and quantifier presented in a 2 (format: verbal or numerical) $\times$ 2 (nutrient type: positive or negative) $\times$ 3 (quantity: low/20%, medium/40%, or high/70%) within-subjects design with four trials per condition. We found that participants looked longer at verbal than numerical quantifiers, and also longer at attribute (nutrient) information with verbal quantifiers, though they did not take longer to make their judgement. Verbal labels also led to more polarised judgements: participants judged positive labels as healthier and negative labels as less healthy with verbal than numerical quantifiers. These judgements were explained by increased attention to the nutrient for low quantities. We discuss how the results fit within theories of information and language processing.

5.2 Introduction

Imagine that an individual is considering whether to buy a cereal bar that is 'high in protein'. Would the cereal bar be more appealing if it stated that it had 70% of protein instead? Varying the format of a quantifier by presenting it in numerical format (i.e., '70%') or verbal format (i.e., 'high') may alter people's judgement and decision-making (Windschitl & Wells, 1996). People's response patterns showed that with verbal quantifiers, their decisions are more influenced by the attribute described ('protein', in the above example) than with numerical quantifiers (see Chapter 4). Some evidence also suggested that compared to verbal quantifiers, decisions with numerical quantifiers rely more on the actual quantity (González-Vallejo et al., 1994). However, this evidence has been based on behavioural outcomes, which provide only indirect evidence of the processes that produce them (e.g., González-Vallejo et al., 1994; Windschitl & Wells, 1996). Without more direct measures of the judgement or decision-making, it is difficult to draw conclusions about the nature of participants' cognitive processing: for example, if they are making a more intuitive decision using verbal quantifiers. The aim of this study was therefore to investigate how people process verbal and numerical quantity information using eye-tracking methodology.

5.2.1 Evidence for differences between verbal and numerical quantifiers

People's behavioural responses differ when they make judgements and decisions based on verbal or numerical quantifiers, which suggests that they approach the judgement or decision differently with the two formats. For instance, people give higher preference ratings on a verbal as opposed to a numerical rating scale (Nicolas et al., 2010), rank products differently when given verbal or numerical scales (Maciejovsky & Budescu, 2013), and believe chance events to occur at higher frequencies when given verbal rather than numerical likelihoods, even when they see the same objective probability (Windschitl & Wells, 1996). Further, when people are told the winning likelihood of gambles, the amounts they bid on the gambles vary more when given verbal than numerical probabilities (Budescu

& Wallsten, 1990).

One might think that these observed differences in judgement and decisions are simply due to the vagueness surrounding the meanings of verbal quantifiers, whereby a verbal quantifier can mean different numerical values to different individuals (Budescu & Wallsten, 1995). However, this presumes that people spontaneously translate verbal quantifiers into numerical values (or vice versa), which may not be the case: they may instead form a general mental representation of the quantity (Reyna & Brainerd, 1991). Further, if vagueness explained variations in judgement and decision outcomes, one would expect greater variability around mean judgements with verbal quantifiers compared to numerical ones, but not a difference between the means (for example, one would observe a greater standard deviation around similar mean ratings). Vagueness cannot explain why mean judgements and decisions should be consistently higher for verbal than numerical quantifiers (Nicolas et al., 2010). Nor can it explain why people are more susceptible to judgement biases with verbal than numerical quantifiers (Windschitl & Wells, 1996). Vagueness also does not explain why people's decisions take into account how positive or negative the attribute information is when given verbal vs. numerical quantifiers (see Chapter 4). Instead, these findings suggest more than a difference in translation: people are making their judgements and decisions in different ways for the two formats.

One explanation for differences in how people make judgements and decisions with verbal and numerical quantifiers is that people process verbal quantifiers more intuitively than numerical ones. This hypothesis is borne out by evidence that compared to numerical quantifiers, verbal quantifiers resulted in people displaying greater levels of cognitive biases associated with quick, intuitive decision-making (Windschitl & Wells, 1996). Dual-process theory defines intuitive processes as automatic, requiring little cognitive effort (see Evans, 2008, and De Neys, 2017b, for overviews). This should result in behavioural markers such as quicker processing and the use of more decision shortcuts (e.g., heuristics; Kahneman, 2011) that can lead to judgement biases (Morewedge & Kahneman, 2010) with verbal than numerical quantifiers. For example, one might reach a decision based on a gut positive feeling rather than a thorough evaluation of all the information (Shiv & Fedorikhin, 1999; Slovic et al., 2007).

Although there is a tacit agreement that verbal quantifiers are more natural (Wallsten et al., 1993), which should indicate that they are more intuitively processed, empirical evidence for the different behavioural markers of intuition is not unanimous. Attempts to trace verbal vs. numerical quantifier processing through reaction time measures showed inconsistent results: studies have found faster decisions with verbal quantifiers (Budescu & Wallsten, 1990; Viswanathan & Childers, 1996), but also faster decisions with numerical quantifiers (Jaffe-Katz et al., 1989; Viswanathan & Narayanan, 1994), and still also no overall difference between the formats (González-Vallejo et al., 1994). Applying a memory load to decisions with verbal or numerical quantifiers (which should constrain the ability to make analytical, but not intuitive, decisions) also showed no impact on decisions for either format (see Chapter 4). On the other hand, past work also found evidence for more use of biases typical of decision shortcuts with verbal than numerical quantifiers (Welkenhuysen et al., 2001; Windschitl & Wells, 1996). For example, participants described the verbal probability of winning a lottery with ten in a hundred balls as higher than one with one in ten balls, but they were less prone to this bias with numerical probabilities (Windschitl & Wells, 1996).

The conflicting evidence from the literature indicates that there is a more complex processing difference between verbal and numerical quantifiers, which leads to overall similarities in certain measures (e.g., response times), but different paths to reach a judgement or decision. For example, although (González-Vallejo et al., 1994) found that on average, people chose similar options in games of chance with verbal and numerical quantifiers, their participants tended to pick gambles with numerical quantifiers that had higher probabilities, and gambles with verbal quantifiers that had higher outcome values (i.e., larger pay-outs). This suggests that the information considered by the participants was different based on format, although it is hard to determine whether these strategies reflect shortcuts based on the quantifier format because both probabilities and outcomes were numerical values. In our own previous work, we found similar patterns in a decision-making task using quantities of nutrients (see Chapters 3 and 4). In these studies, participants decided whether either a verbal or numerical quantity of a nutrient would sum with a previously presented quantity to exceed a healthy consumption limit. The results found that in general, participants tended to rely more on the valence of the nutrient to make their decision with verbal than numerical quantifiers (e.g., quantities of minerals were more often mistaken as healthy). This was especially the case for a positive nutrient (minerals), but findings for negative nutrients (e.g., fat) were less clear-cut: in some instances, numerical quantities of fat were mistaken as unhealthy more than verbal ones (see Chapter 3). These findings suggest that the format influences one's decision strategy: verbal quantifiers place more importance on positive contextual information (e.g., good attributes and outcomes), and numerical quantifiers on negative information, or actual quantities. However, this conclusion is based on indirect reasoning from the outcomes of decisions, which do not provide direct evidence about which pieces of information are actually attended to and processed. A process-tracing approach could thus supplement outcome measures to test the explanation that the format of a quantifier influences what is important to one's decision-making.

5.2.2 Eye movements as a process-tracing measure for judgement and decision-making

Differences in how people make decisions with verbal and numerical quantifiers should be reflected in attention to information. If verbal quantifiers are more intuitively processed, they should require less attention to process than numerical quantifiers. Similarly, if verbal quantifiers encourage people to rely more on positive contextual information, we would expect more attention to this information with verbal than numerical quantifiers. Some evidence from change detection paradigms suggest this would be the case: Moxey (2017) found that readers who saw two quantified phrases with a minor change to the wording were more likely to notice a change to the attribute in a phrase (e.g., changing 'low fat' to 'low sugar') when the quantifier was verbal than numerical; conversely, readers noticed changes.in the quantity (e.g., changing '5% fat' to '15% fat') when it was presented numerically than verbally. Building on this, differences in attention can be reflected in the pattern of an individual's eye movements across visual information (Russo, 2011).

Researchers are still debating the extent to which eye fixations are an indicator of cognitive processing (Orquin & Loose, 2013; Schulte-Mecklenbeck et al., 2017b): some argue that attention, as indicated by eye movement patterns, does not necessarily mean that the information is being processed, but in general, studies find that longer fixation durations correspond with greater and more costly cognitive processing (Horstmann et al., 2009; Orquin & Loose, 2013). More crucially, it is unlikely that information is used if it is not attended to (Siegrist et al., 2015). Tracking what information people attend to via the pattern of their eye movements has become an increasingly applied method to inform how information is processed and feeds into judgements across many decision-making domains (Orquin & Holmqvist, 2018; Schulte-Mecklenbeck et al., 2017a). Therefore, eye-tracking provides a measure to trace the processes of judgements and decisions with verbal and numerical quantifiers.

5.2.3 The current study

The goal of this study was to investigate visual attention to attribute and quantity information when the presented quantifier was either verbal or numerical, and how this affects judgement of the information. To do so, we used the context of nutrition labelling, where we could vary both the nutrient (the 'attribute') and the quantifier presented. We built on the postulate that verbal quantifiers are more intuitively processed than numerical quantifiers, and previous empirical work that found decision outcomes to show different strategies for verbal and numerical quantifiers.

First, we hypothesised that people would pay more attention to (i.e., gaze longer at) the quantifier when it was numerical than verbal. It is worth noting here that attention times to different quantifier formats may also be affected by basic differences in stimuli length or character types (i.e., numbers vs. letters; Orquin & Holmqvist, 2017). For example, a longer word could result in longer and more fixations. To control for these differences, we developed verbal and numerical labels similar in length, and conducted robustness checks that statistically controlled for quantifier length.

Second, we expected that people's judgements would follow patterns found in our previous work (see Chapter 4): positive (negative) nutrients would be rated more (less) healthily with verbal than numerical quantifiers. We hypothesised that different patterns of attention would explain these judgement differences. We tested whether people would pay more attention to positive nutrients with verbal than numerical quantifiers, and more attention to the negative nutrient with numerical than verbal quantifiers. In line with recent scientific guidelines, the hypotheses, methods, and statistical analyses were registered prior to conducting the experiment. These are available along with experimental data on the Open Science Framework.

5.3 Method

5.3.1 Participants

The study was powered to detect an effect size of f = .10 (α= .05, 1-β= .80, two-tailed test)[1]. Participants were 149 students who completed the study for course credit (78% female; age range 18-46 years, $M = 20.5$, $SD = 4.7$). Data from one participant was excluded due to a programming glitch during their session.

5.3.2 Design

We tracked participants' visual attention as they judged 48 nutrient labels, each with a single quantity and a single nutrient (see Figure 5.1 for an illustration). The 48 labels resulted from the crossing of three variables in a within-subjects design with four trials per condition: 2 (format: verbal or numerical) $\times$ 2 (nutrient type: positive [protein or minerals] or negative [saturated fat or sugar]) $\times$ 3 (quantity: low/20%, medium/40%, or high/70%).

[1] We planned to conduct multi-level analyses (MLM) to take into account data from individual trials clustered within a participant, but we powered the study based on a three-way ANOVA as this allowed us to determine a sample size estimate for a small format effect without knowing in advance the beta parameters for the fixed and random effects that would be necessary to fit the MLM.

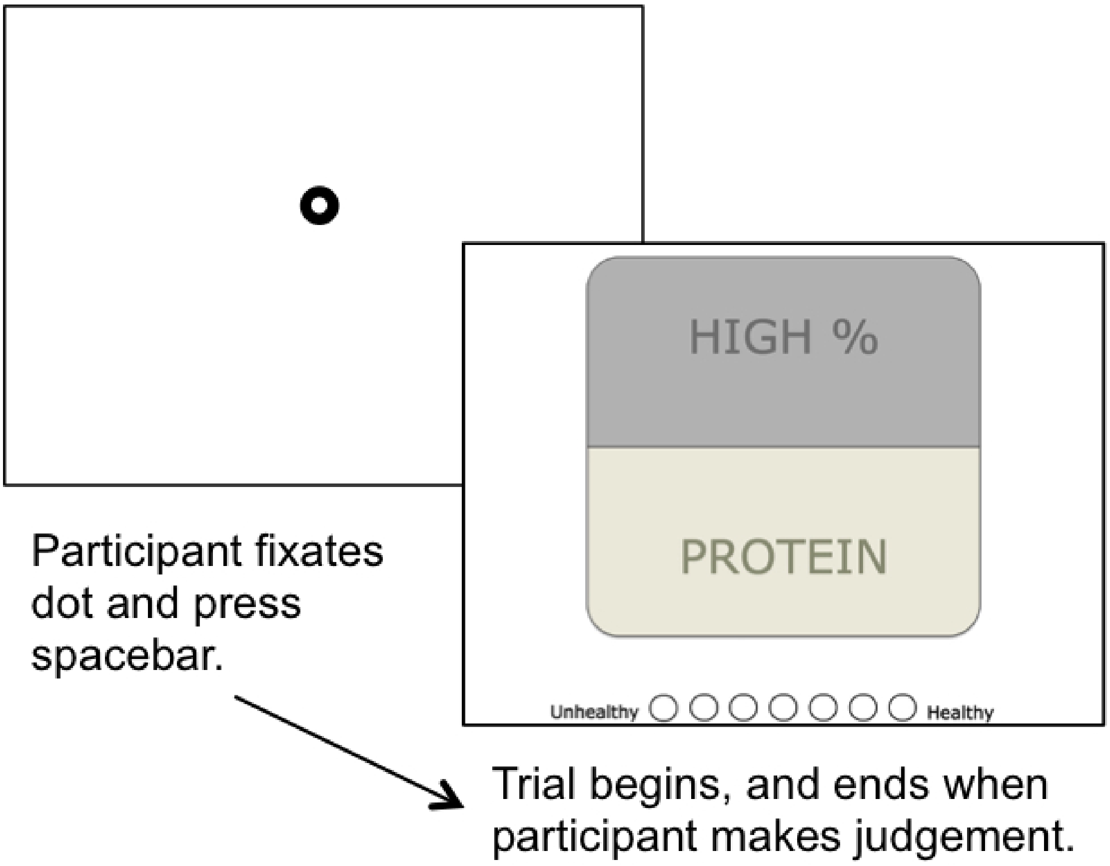

Figure 5.1. Procedure for one trial in the experiment, showing an example of a high protein label. The fixation dot was aligned to be in the centre of the nutrient label when it appeared.

Note. Examples of how the labels and response scale were counterbalanced are provided in Appendix D.

5.3.3 Materials

We designed simplified nutrition labels (see Figure 5.1 for an example of a verbal label) that featured the quantity of a nutrient in terms of its percentage contribution to one's Guideline Daily Amount (GDA; i.e., the total amount one should eat in a day). For half the participants, the nutrient was at the top of the label, with the quantifier below, and for the other half, it was the opposite (see Appendix D for an example of a numerical label in the four counterbalanced conditions). We measured participants' eye fixations within two Areas of Interest (AOIs): the quantifier and the nutrient portion of the label. Each AOI was the same size and subtended approximately 19 by 10 degrees of visual angle.

To derive comparable quantities for the verbal and numerical quantifiers, we selected numerical quantities that are perceived on average psychologically equivalent to the verbal quantifiers used (low, medium, and high %) according to a previous translation study on a similar participant sample (see Chapter 2). This method is typically used to compare verbal and numerical quantifier judgements (Teigen & Brun, 2000; Welkenhuysen et al., 2001). Although we considered the possibility of using a verbal-numerical pairing provided by national nutritional guidelines (e.g., UK Department of Health, 2016), prior research shows that people do not interpret verbal quantifiers as the standard indicates (see Chapter 2); we would thus expect participants' psychological interpretations to widely differ from official translations (as is the case for probabilities, Budescu et al., 2009, 2012, 2014, and frequency quantifiers, Berry et al., 2002, 2003, 2004; Carrigan et al., 2008; Hamrosi et al., 2012; Knapp et al., 2015).

In addition, to increase the comparability of the quantifier lengths, we reduced 'MEDIUM %' to 'MED %'. After this modification, all the quantifiers were between three and five characters long (e.g., LOW % had four characters while 20 % had three characters).

5.3.4 Procedure

Participants were tested individually in the laboratory. Upon arrival, participants signed a consent form outlining the experimental procedure. Participants' heads were stabilised on a chin rest to restrict movement. We used an EyeLink1000 eye-tracker (SR Research, 2007) mounted on the desk below a 17-inch PC monitor (screen resolution 1024×768) to track pupil image and corneal reflection of participants' right eye. The monitor was approximately 60cm from the corneal surface, with dim background lighting. The EyeLink1000 has a sampling rate of 1,000 Hz, average accuracy of 0.25-0.5°, and spatial resolution of 0.01°. The experimenter performed a 9-point calibration check prior to starting the experiment.

The task was presented using SR Research's Experiment Builder software (programme available on the OSF). Participants read instructions about the task on the screen, which stated that their goal was to evaluate how healthy a food with each label was. The instructions also included an example of what the labels looked like, and what a GDA of a nutrient meant. Participants also read a definition of each nutrient they would see during the experiment (e.g., '*Sugar refers to any of several sweet carbohydrate substances*'). In each trial, participants viewed a label and assessed how healthy they thought the food with the presented label was. Participants used a mouse to give their judgement on a 7-point Likert scale below the label (unhealthy to healthy or healthy to unhealthy; randomly assigned between participants). Each trial started with a fixation dot appearing on the screen between the position of the quantifier and attribute interest areas to ensure participants' gaze would fall on the centre of the nutrition label at the start of each trial (see Figure 5.1). Participants fixated the dot and pressed the spacebar to begin the trial. The trial ended once they made their judgement.

Participants first completed a practice set of six trials and had the opportunity to ask questions before beginning the experimental trials. Participants performed two blocks of twenty-four trials with a break in between. The experimenter performed another calibration check before continuing with the second block if the participant moved their head during the break. In the first block, participants were randomly assigned to view either verbal (e.g., 'low %') or numerical (e.g., '20 %') percentages. In the second block, the quantifiers were in the other format. Within a block (randomly presented), the nutrients were either positive (minerals or protein) or negative (saturated fat: 'sat fat', or sugar), and the quantifiers were low/20%, med/40%, or high/70%.

At the end of the experimental task, participants completed a questionnaire that assessed socio-demographic information.

5.4 Variables and Analysis Strategy

For each trial, we measured the following variables: fixation duration and number of fixations on the quantifier and nutrient AOIs, and healthiness judgement. Fixations were determined according to the standard EyeLink algorithm in cognitive configuration. This detects saccades whenever the eye exceeds velocity, acceleration, and motion thresholds of 30°/sec, 8000°/sec, and 0.15°/sec respectively (SR Research, 2007). This resulted in a minimum threshold of 80ms to define a fixation. Total fixation duration was defined as the sum duration of all fixations on the AOI, which represents the total amount of attention paid to this information during a trial.

We tested our hypotheses in a pre-registered multilevel model using SPSS, after excluding data from nine (0.13%) trials where 0 fixations and 0ms fixation duration was recorded, as this indicated that participants' decisions were made without looking at the label, and their attention patterns could thus not be informative in these trials. We used a variance components model and included fixed effects for format, nutrient type, and quantity, and their interactions. The full random effects model did not converge, hence we dropped individual slopes systematically until a convergent model was obtained, which included only by-participant intercepts.

To further test the extent to which attention to the nutrient was responsible for explaining healthiness judgements, we performed pre-registered secondary mediation analyses using the PROCESS macro for SPSS (Model 5; Hayes, 2013). Because nutrient type moderated the effect of format on healthiness judgements, but not fixations, we included nutrient type as a moderator of the direct effect between format and judgement, but not in the mediated pathway. The analyses used bootstrapped confidence intervals based on 5,000 samples to investigate the effect of format on judgement as mediated by total fixation duration on the nutrient (in seconds) for each of the three quantities, while controlling for the moderating effect of nutrient type in the direct relationship. Figure 5.2 illustrates the mediation model.

5.5 Results

5.5.1 General patterns in attention and judgement

We provide first an overall description of the eye movement and behavioural response data. Participants spent on average 2.94s ($SD = 1.67$) fixating, and made 11.32 ($SD = 6.43$) fixations in a trial. Participants looked less at the

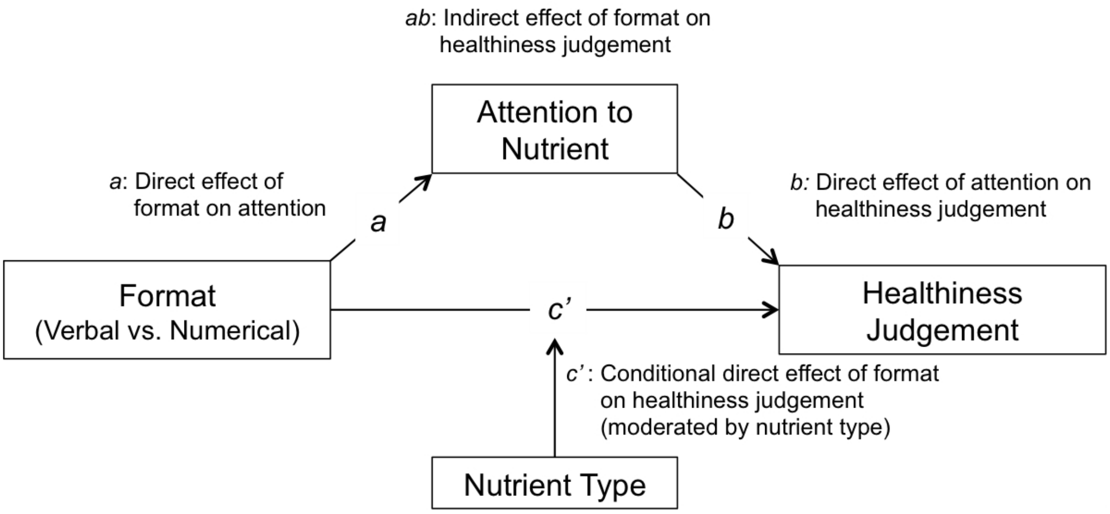

Figure 5.2. Mediation model for the effect of format on healthiness judgements at each of the three quantities, specifying the direct and indirect effects. Values for the beta coefficients and their 95% confidence intervals are given in Table 5.2.

Note. Only the mediation pathway for low/20% quantities was significant, indicating that paying more attention to the nutrient led to higher healthiness judgements for low (vs. 20%) of negative nutrients and lower healthiness judgements for low (vs. 20%) of positive nutrients.

quantifier (M_{duration} = 594ms, SD = 607; M_{number} = 2.66, SD = 2.30) than the nutrient (M_{duration} = 710ms; SD = 721; M_{number} = 3.04, SD = 2.56). Figure 5.3 illustrates the overall fixation counts across different areas of the stimulus screen for verbal and numerical quantifiers, and Figure 5.4 provides an example of scan paths for a participant for different trial conditions, illustrating the longer attention patterns for attributes with verbal than numerical quantifiers. The number and duration of fixations were highly correlated, $r = .86$, $p < .001$ (nutrient), $r = .88$, $p < .001$ (quantifier), and the pattern of results for number of fixations was similar to the results for total fixation duration. Results of further analyses on number of fixations are reported in Appendix D (Table D.1).

We checked if overall fixation duration and response times differed between formats; neither was significant, $t(147) = 1.49$, $p = .139$ (fixation duration); $t(147) = 0.97$, $p = .333$ (response time). Participants took on average 3.70s to make their judgements (SD = 2.13), and rated the labels on average 3.94 (SD =

1.93) on the 7-point Likert scale.

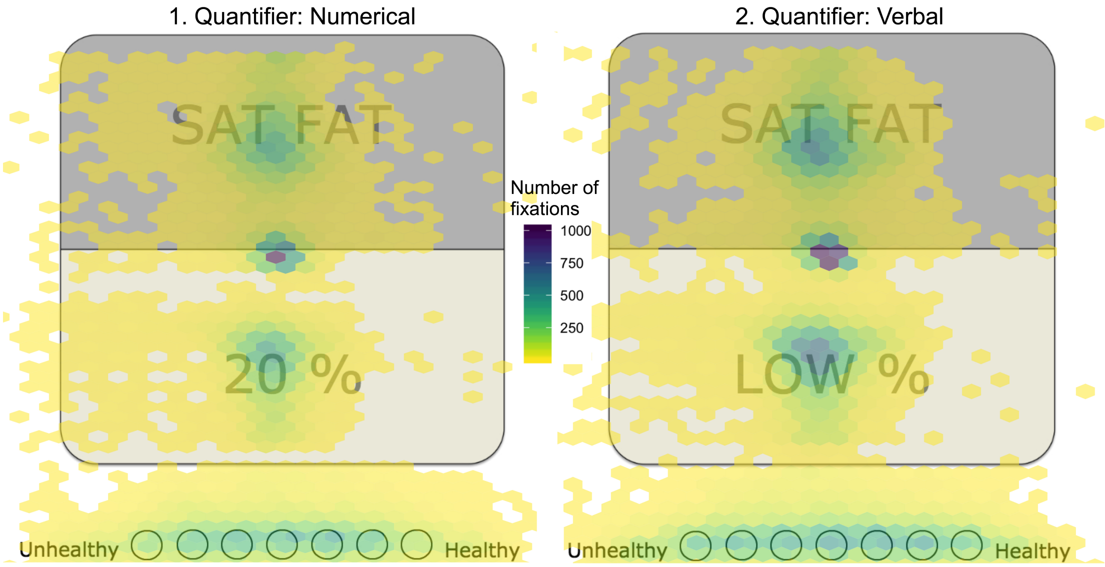

Figure 5.3. Fixation density plot illustrating the combined number of fixations on areas of the screen across participants for trials with numerical labels (left) vs. verbal (right). Darker colouring indicates a greater number of fixations.

5.5.2 Results of multilevel analysis

The results for total fixation duration on nutrient and quantifier AOIs, and healthiness judgement are reported in Table 5.1. We address below the tests of each of our hypotheses.

Visual attention to quantifiers. Based on our first hypothesis, we would expect a main effect of format on fixations on the quantifier, with fixations being greater for the numerical than verbal quantifiers. Contrary to this, participants looked longer at verbal than numerical quantifiers (M_{verbal} = 638ms, *SD* = 571; $M_{\text{numerical}}$ = 551; *SD* = 638), $F(1, 7083) = 45.98$, $p < .001$. As fixation duration was correlated with the length of the quantifier ($r = .05$), we conducted additional checks to ascertain if this was an artefact of the typically longer verbal quantifier lengths compared to numerical. The effect was still significant when we tested format as a fixed effect while controlling for character length, $F(1, 7092) = 35.52$, $p < .001$. However, running the multilevel analysis on total fixation duration per quantifier character (Saikh et al., 2015) found that duration per character was longer for numerical than verbal quantifiers (M_{verbal} = 149ms per character,

Table 5.1. Fixed and interaction effects for format, nutrient type, and quantity in the pre-registered multilevel analyses for total fixation duration on the nutrient AOI, total fixation duration on the quantifier AOI, and healthiness judgements. Effects specific to our hypotheses are bolded and indicated with #.

Factor	***F*** **(error df = 7083)**	**df**	***p***
Total fixation duration on quantifier AOI			
# Format (verbal or numerical)	**45.98**	**1**	**< .001**
Nutrient type	0.33	1	.565
Quantity	42.74	2	< .001
Format × nutrient type	2.97	1	.085
Format × quantity	6.62	2	.001
Nutrient type × quantity	1.39	2	.250
Format × nutrient type × quantity	1.42	2	.242
Total fixation duration on nutrient AOI			
Format (verbal or numerical)	5.95	1	.015
Nutrient type (positive or negative)	7.46	1	.006
Quantity (high, med, or low)	10.23	2	< .001
# Format × nutrient type	**1.09**	**1**	**.296**
Format × quantity	1.61	2	.201
Nutrient type × quantity	2.83	2	.059
Format × nutrient type × quantity	1.38	2	.251
Healthiness judgements			
Format (verbal or numerical)	44.07	1	< .001
Nutrient type (positive or negative)	2542.35	1	< .001
Quantity (high, med, or low)	77.87	2	< .001
# Format × nutrient type	**193.25**	**1**	**< .001**
Format ×s quantity	11.70	2	< .001
Nutrient type × quantity	3828.34	2	< .001
Format × nutrient type × quantity	198.35	2	< .001

Note. Results of multilevel analyses for number of fixations are reported in Appendix D.

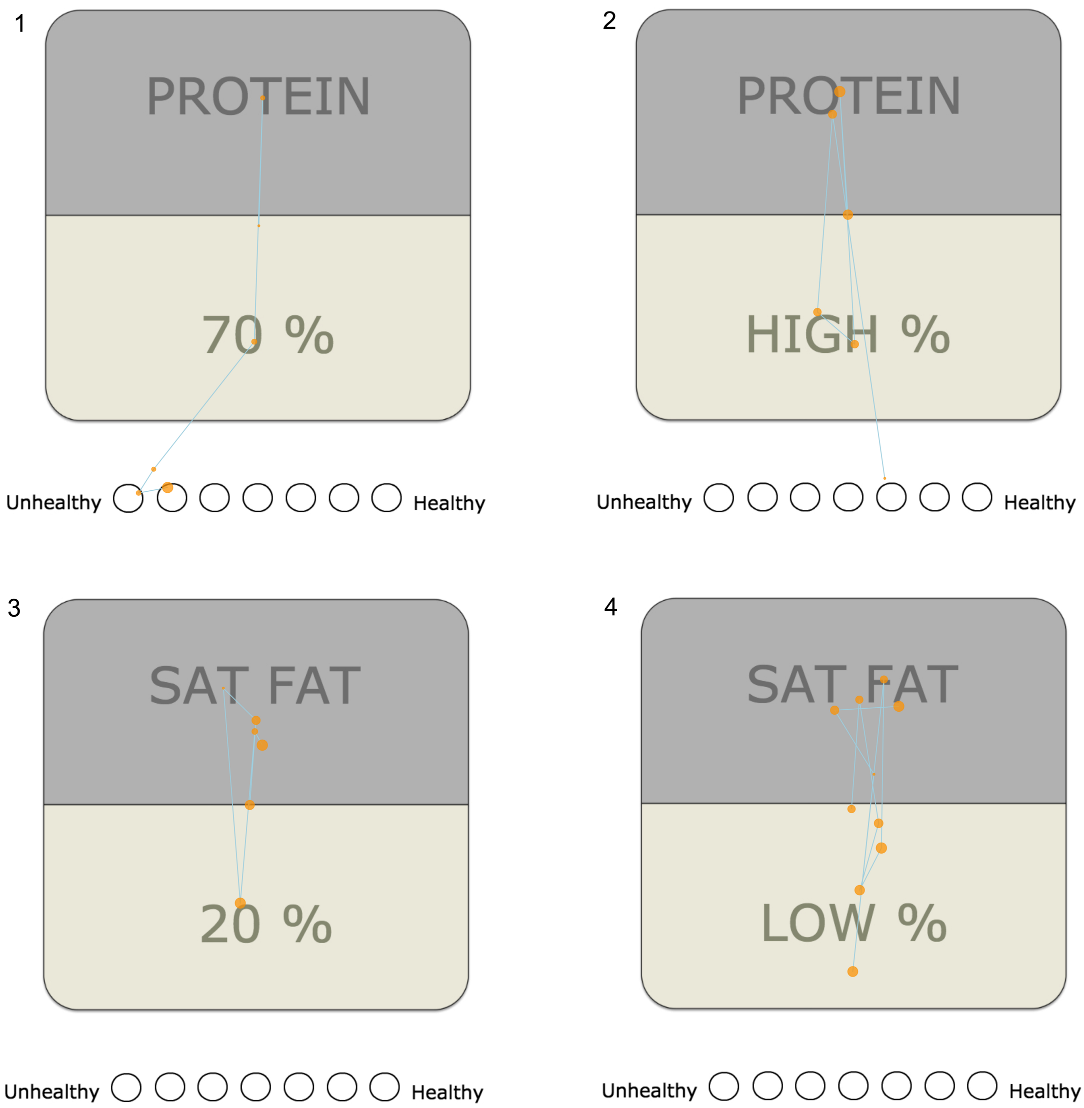

Figure 5.4. Examples of scan paths (blue lines) of fixations (in orange) for one participant from each of the trial conditions: (1) numerical-positive, (2) verbal-positive, (3) numerical-negative, and (4) verbal-negative label. Larger dots correspond to longer fixations.

SD = 138; $M_{\text{numerical}}$ = 184ms per character, SD = 213), $F(1, 7083) = 80.01$, $p < .001$. Due to this discrepancy, we conducted a further check using a linear regression analysis with cluster-corrected standard errors, using the R package 'miceadds' (Robitzsch et al., 2019). This analysis did not find a significant effect of format, $b = 72.10$, $p = .068$.

Visual attention to nutrients. Our hypothesis regarding attention to nutrients predicted an interaction effect between format and nutrient type for

Table 5.2. Pairwise comparisons for attention to nutrient between verbal and numerical labels

Format	Nutrient type	Mean fixation duration (ms)	*SD*
Verbal	Positive	740	742
	Negative	716	632
	Overall	728	609
Numerical	Positive	720	797
	Negative	665	699
	Overall	693	750

Note. Mean difference between format was significant overall and for negative nutrients. The interaction between format and nutrient type was not significant.

fixations on the nutrient, with fixations being greater for the verbal than numerical quantifiers on positive nutrients, but the opposite for negative nutrients. We did not find this interaction, $F(1,\ 7083) = 1.09$, $p = .296$. In contrast, our planned pairwise comparisons between quantifier formats for each nutrient type found that people looked longer at both types of nutrients with verbal than numerical quantifiers, but this was only significant for negative nutrients, $F(1,\ 7083) = 0.97$, $p = .324$ (positive); $F(1,\ 7083) = 6.06$, $p = .014$ (negative). The means and standard deviations for fixation duration in each condition is given in Table 5.2. Overall, the greater attention to the nutrient for verbal than numerical quantifiers was significant $F(1,\ 7083) = 5.95$, $p = .015$.

Healthiness judgements. Based on our second hypothesis, we expected an interaction effect between format and nutrient type for healthiness judgements, with healthier judgements for verbal than numerical quantifiers for positive nutrients, but the opposite for negative nutrients. As expected, we found a significant two-way interaction between format and nutrient type, and this was further explained by the significant three-way interaction between format, nutrient type, and quantity. As shown in Figure 5.5, participants were sensitive to how quantities modify the overall valence of a label: they judged smaller amounts of negative nutrients as healthier than positive ones, but larger amounts of positive nutrients as healthier than negative ones, $F(2, 7083) = 198.35$, $p < .001$. Participants rated labels with overall positive valence (e.g., 'low fat' or 'high minerals') as healthier in verbal than numerical format, but labels with overall negative valence (e.g., 'low minerals' or 'high fat') as healthier in numerical than verbal format. This valence effect was larger for the verbal than numerical quantifiers for low/20% and high/70% quantities, but the opposite was observed for med/40% quantities. Pairwise comparisons between the formats for each quantity-nutrient type combination are provided in Table 5.3.

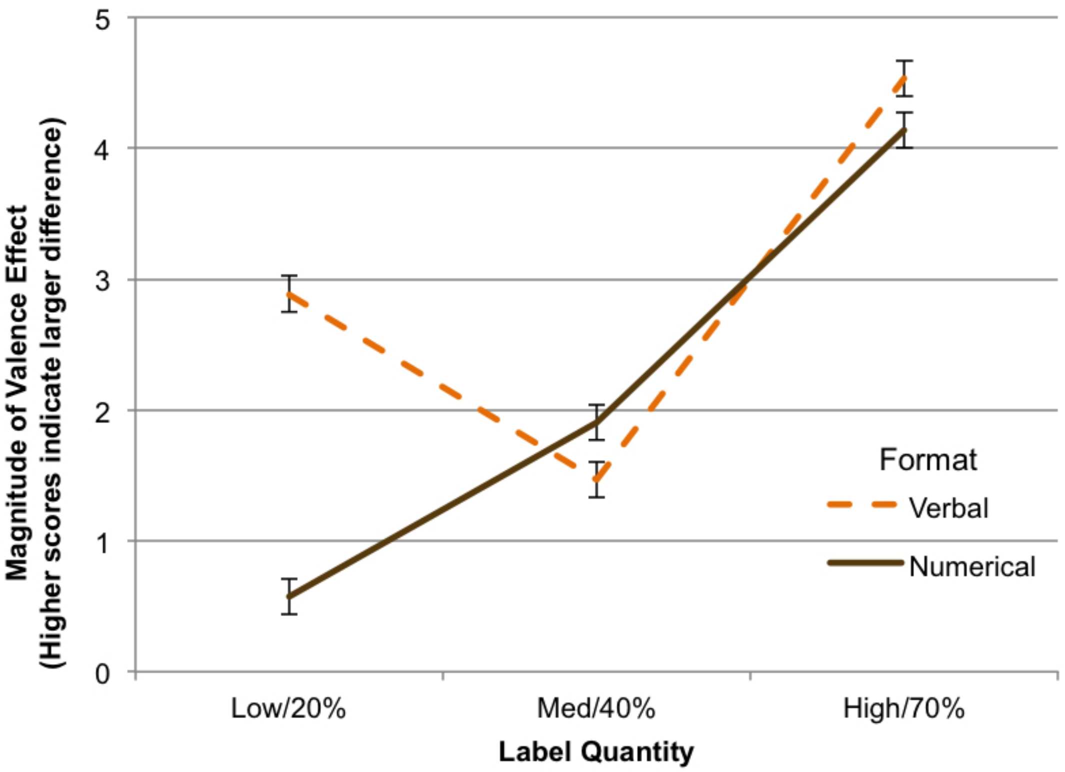

Figure 5.5. Valence effect between formats at low (20%), med (40%), and high (70%) quantities. Higher scores indicate a greater difference between positively and negatively valenced labels (e.g., 20% fat vs. 20% minerals). Error bars reflect 95% confidence intervals.

Table 5.3. Pairwise comparisons for healthiness judgements between verbal and numerical quantifiers.

Quantity	Nutrient type	Mean difference in healthiness (verbal –numerical)	95% CI	$F(1, 7083)$	p
Low/20%	Positive	-1.03	[-1.16, -0.89]	217.71	< .001
	Negative	1.29	[1.15, 1.42]	343.35	< .001
Med/40%	Positive	0.16	[0.02, 0.29]	5.03	.025
	Negative	0.60	[0.46, 0.74]	74.29	< .001
High/70%	Positive	0.25	[0.12, 0.39]	13.13	< .001
	Negative	-0.14	[-0.28, -0.002]	3.96	.047

Note. Healthiness judgements were made on a 1-7 Likert scale.

5.5.3 Results of the pre-registered secondary mediation analyses: Visual attention mediated healthiness judgements only for small quantities

The results of the moderated mediation analyses are shown in Table 5.4. We found a significant indirect effect of format on healthiness judgements through participants' attention to the nutrient (total fixation duration in seconds) for low/20% quantities only, b_{low} = .006, CI [.001, .016]; b_{med} < .001, CI [-.001, .006]; b_{high} = .001, CI [-.004, .010]. This indicated that for small quantities, the increase in attention paid to the nutrient with verbal quantifiers (compared to numerical) led to participants judging negative nutrients less healthily and positive nutrients more healthily.

5.6 Discussion

The goal of this study was to address the lack of direct evidence for patterns of judgement and decision-making with verbal and numerical quantifiers. The literature has posited that verbal quantifiers are more intuitively processed (Windschitl & Wells, 1996), but behavioural results do not consistently support

Table 5.4. Direct and indirect effects (via attention) of format (verbal vs. numerical) on healthiness judgements found in the moderated mediation analyses for low/20%, med/40%, and high/70% quantities.

Effect pathway	**Quantity**	**Beta coefficient**	**95% CI**
Effect of format on attention to nutrient (a)	Low/20%	-0.071	[-0.129, -0.013]
	Med/40%	-0.020	[-0.082, 0.042]
	High/70%	-0.013	[-0.067, 0.041]
Effect of attention to nutrient on healthiness (b)	Low/20%	-0.089	[-0.149, -0.019]
	Med/40%	-0.019	[-0.099, 0.061]
	High/70%	-0.107	[-0.198, -0.017]
Indirect effect of format on healthiness through attention (ab)	Low/20%	0.006	[0.001, 0.017]
	Med/40%	-0.020	[-0.001, 0.006]
	High/70%	-0.013	[-0.004, 0.009]
Direct effect of format on healthiness (c')	Low/20%	2.314	[2.082, 2.545]
	Med/40%	0.443	[0.266, 0.620]
	High/70%	-0.381	[-0.559, -0.202]

Note. A negative coefficient for format indicates lower attention for the numerical as compared to verbal condition.

this, and decision outcomes indicate that people may use different decision strategies with verbal and numerical quantifiers that predict attention to different information (González-Vallejo et al., 1994). We used eye-tracking to determine the level of attention given to attribute and quantity information when the quantifier was verbal or numerical, and assess the influence of attention on judgement.

We derived three hypotheses based on previous work. First, we expected that people would look longer at numerical quantifiers, which could indicate more attention and processing effort than for verbal quantifiers. Second, we expected participants' judgements to be healthier for positive nutrients and less healthy for negative nutrients. Third, we expected participants to use different information for judgements with verbal and numerical quantifiers, as indicated by more attention to a positive nutrient for verbal than numerical quantifiers, and vice versa for negative nutrients. Not all our hypotheses were supported. Participants paid more attention to verbal than numerical quantifiers. Participants' judgements were indeed more polarised with verbal and numerical quantifiers, but this effect took into account the overall valence of the nutrient *and* its amount. Finally, we found that the verbal format increased attention to nutrients, both positive and negative, rather than in different directions for each. The effect of quantifier format on healthiness judgement was mediated by attention to the nutrient, but only for small quantities.

5.6.1 People pay attention to different components of information with verbal and numerical quantifiers

Attention to quantifiers. Contrary to what we expected, participants looked longer at verbal than numerical quantifiers, indicating that people paid more attention to verbal quantifiers. This pattern of results should be taken with caution as the verbal labels were slightly longer, and controlling for the number of characters led to a different result: participants paid more attention per character to numerical than verbal quantifiers. Further, when using a different statistical technique, we found no significant effect. Due to the ambiguity in interpreting the longer fixation durations for verbal quantifiers, the results do not provide conclusive support for more intuitive processing of verbal quantifiers. One could interpret that greater attention per character reflects an ability to identify words as a unit rather than by reading each letter (Healy, 1976) —which should not be the case for numbers (Orquin & Loose, 2013). This interpretation would suggest that verbal quantifiers are more intuitive. However, if participants were identifying words quicker, longer overall fixations would still indicate that participants paid more attention than simply reading the verbal quantifier should require, even if they had a greater length than the numerical ones. This interpretation would suggest that verbal quantifiers are *less* intuitive. Given that participants did not spend significantly more time looking at the screen or making their judgements for either quantifier, there is less support that the quantifiers differ in terms of the level of intuitive processing involved.

Attention to nutrients. The ability to distinguish attentional processes involved in reading the nutrient when the quantifier is verbal or numerical is less debatable, because this attribute remains constant across conditions. Building on the work in Chapters 3 and 4 of this thesis, which showed that people use different decision strategies with different quantifier formats (participants relied more on positively valenced nutrients for verbal quantifiers and negatively valenced ones for numerical quantifiers), we expected that participants would pay greater attention to positive or negative nutrients with verbal and numerical quantifiers respectively. Instead, we found that participants paid more attention paid to the nutrient for verbal than numerical quantifiers, whether the nutrient was positive or negative. This suggests that the very same attribute, such as the word 'SUGAR', could be processed for longer when paired with verbal than numerical quantifiers.

The unexpected finding that people paid greater attention to nutrients paired with verbal than numerical quantifiers suggests that rather than the use of different decision shortcuts between the formats (e.g., relying on either positive or negative nutrients), people used shortcuts more with verbal than numerical quantifiers. The more polarised judgement outcomes with verbal quantifiers support this interpretation that people relied more on the nutrient to make their judgements in this condition, and is also consistent with past work suggesting people use such strategies more with verbal than numerical quantifiers (González-Vallejo et al., 1994; Windschitl & Wells, 1996). To determine if attention on a nutrient had an effect on judgement, we ran mediation analyses. These analyses found that attention mediated the effect of format on healthiness judgements for low (20%) but not medium (40%) or high (70%) quantities.

5.6.2 Explaining attention and judgement patterns for verbal and numerical quantifiers

One could argue that participants judged large positive verbal quantities to be more healthy and small negative verbal quantities to be less healthy than the corresponding numerical ones because they translated the verbal quantifiers to mean less than the numerical ones (e.g., in line with official guidelines that translate high, for example, as around 30%; Council of the European Union, 2006). We believe this is unlikely for two reasons. First, it is well-documented that people overestimate official guidelines for translating verbal quantifiers (Berry, 2006; Budescu et al., 2012; Carrigan et al., 2008; Knapp et al., 2010, 2009b); we also found this to be the case for nutritional guidelines (see Chapter 2). Second, our data showed that large verbal quantities (e.g., high protein) were unlikely to be evaluated as less than the corresponding numerical quantity (i.e., 70% protein): for example, a high protein label was healthier than a 70% protein label, which fits more with an intuitive understanding than an underestimation of 'high'. Finally, the attention patterns to the nutrient —greater for verbal than numerical quantifiers —provides an additional indication that participants considered this information more with verbal quantifiers.

A greater reliance on nutrient information for verbal quantifiers could be explained in two ways. First, people might be using the nutrient's valence (positive or negative) to make a more intuitive judgement. However, because participants did not display other indicators of intuitive processing (quicker responses, less attention to the quantifier) for verbal more than numerical quantifiers, this explanation is less plausible. The second explanation is that people are using the contextual information functionally to guide their judgements (Horn & Ward, 2006). Verbal quantifiers, because they are more natural in language, could act as a signal to encode more contextual information from the attribute, thus resulting in more attention to the nutrient.

The pragmatic properties of verbal quantifiers could also explain why attention to the nutrient only mediated healthiness judgements for the low, but not medium and high, quantifiers. Another pragmatic property of verbal quantifiers is their ability to provide implicit information about what information a reader should focus on (Moxey & Sanford, 1986; Teigen & Brun, 1995). Most quantifiers, like medium and high, are 'positive', which means that they focus on the attribute described (Sanford & Moxey, 2003). For example, '*the meat is high in fat*' puts a focus on the amount of fat. In contrast, some quantifiers, like low, are 'negative', which means that they focus on an alternative attribute not mentioned in the phrase. For example, '*the meat is low in fat*' puts a focus on how lean the meat is. Studies comparing similar verbal and numerical quantities show that while verbal quantifiers have this focusing property (e.g., low is negative), numerical quantifiers are more ambiguous, with small quantities being understood as focusing on either the presence or absence of the attribute (Teigen & Brun, 1995, 2000). Therefore, while verbal quantifiers showed increased attention to the nutrient relative to numerical quantifiers, the subsequent difference in judgement would have been greater for the small quantities because the verbal 'low' had negative directionality but the numerical '20%' could be considered more neutral (Teigen & Brun, 2000). The predictions of pragmatic theory —that attributes should receive more attention with verbal than numerical quantifiers, and this attention would shift to a different attribute between positive and negative quantifiers —present a promising future direction to test for attention differences between verbal and numerical quantifiers.

5.6.3 Implications for food labelling

A practical question following from differences in judgement and attention between verbal and numerical quantifiers is whether greater attention to an attribute will result in a better decision. We found that greater attention to a nutrient led to participants judging a low % of fat or sugar as healthier, but a low % of minerals or protein as less healthy than the numerical quantity. Because the quantities were of 'guideline daily amounts', which are standardised percentages of the daily total one should consume, the quantity expresses the same proportion of recommended consumption whether the nutrient is sugar or protein (Rayner et al., 2004). From this standpoint, the more polarised judgements of verbally-quantified food are irrational.

However, before concluding that verbal quantifiers lead to suboptimal decisions, the goal of the decision-maker should be taken into consideration. In practice, whether a judgement difference is better or worse should depend on the goal of the communication and the structure of the decision environment (Gigerenzer & Todd, 1999; Simon, 1990). Health issues surrounding food consumption in developed nations tend to revolve around overconsumption and its contribution to obesity (World Health Organization, 2016). From this standpoint, a format that encourages reduced consumption of fat and sugar could be beneficial. The judgements for large quantities were also significantly different between verbal and numerical quantifiers, and in the other direction, although we did not find evidence that this was a result of increased attention to the nutrient. This would suggest that verbal formats could encourage reduced consumption of fat and sugar but greater consumption of protein and minerals. One caveat, however, is that we did not test attention patterns when both a verbal and numerical quantifier are paired with different combinations of nutrients in the same label. In cases where more than one nutrient is presented, people tend to attend to negative nutrients more than positive (Graham & Jeffery, 2011; Miller et al., 2015). However, our results suggest a possibility that if one nutrient has a verbal quantifier and the other a numerical quantifier, this could alter attention patterns. This could be beneficial in a case of a cereal with a high amount of fibre and small amounts of sugar, but less so if the high fibre obscures excessive amounts of sugar in the same cereal. As current systems of food labelling allow verbal quantifiers for positive elements to be paired with numerical quantifiers for negative ones, this is a direction worth pursuing in future research.

5.6.4 Conclusion

This work extends previous research by using eye-tracking methodology to show that people's attention processes differ when given verbal or numerical quantifiers. Verbal quantifiers led to greater attention on attribute information, and also more polarised judgements of healthiness for nutrient labels. These attention patterns bear implications for the communication of quantified information. This study used nutrition labels as an example, but virtually all other quantified communications, such as risk information, event forecasting, or news reporting involve both attributes and quantifiers. We suggest that the use of verbal or numerical quantifiers to communicate quantity information should depend on the goals and context of the communication.

Chapter 6: The Directional Focus of Verbal and Numerical Quantifiers Affects the Attribute Framing Effect

6.1 Abstract

People find positive attribute frames (e.g., 75% lean) more persuasive than negative ones (e.g., 25% fat). In three pre-registered experiments, we tested whether this effect would be magnified by verbal quantifiers. This moderating effect of quantifier format was predicted by previous empirical work and two posited properties of verbal quantifiers. First, verbal quantifiers are a more intuitive format than numerical quantifiers, and should predispose people more to judgement biases such as the framing effect. Second, verbal quantifiers have the ability to put focus on attributes, which should provide a strong linguistic signal that the positive frame is better than the negative one. Using a mixed design varying frame (positive or negative) and quantifier format (verbal or numerical) between-subjects, and quantity pairs (e.g., 5% or 25% fat) within-subjects, we found a robust framing effect, but it was not moderated by format. Quantifier format also did not affect whether participants' focus was directed to or away from the cited attribute. When focus was to the cited attribute (e.g., 'lean' in the lean frame), the focus partially mediated the framing effect. These results suggest that focus contributes to the framing effect, but contrary to past work, numerical quantifiers can have similar focusing properties to verbal ones.

6.2 Introduction

The description, or 'frame', that people choose to present an item changes how others judge that item (Tversky & Kahneman, 1981). Logically speaking, it should not matter whether one describes beef as '25% fat' or '75% lean', as these are mathematically equivalent. However, people will judge a 75% lean beef as more desirable than a 25% fat one (Levin & Gaeth, 1988). This 'attribute framing effect', where the positive or negative presentation of an item's attribute affects the evaluation of the item, has been robustly demonstrated across multiple domains, including performance evaluation (Kreiner & Gamliel, 2017; Leong et al., 2017), health decisions (Krishnamurthy et al., 2001), and even mate choice (Saad & Gill, 2014).

Although framing effects —attribute or otherwise[1] —have been widely replicated and studied for decades, questions about what moderates the effect (i.e., factors that increase or decrease its size), and why, remain relevant (Gal & Rucker, 2018; Maule & Villejoubert, 2007). One potential moderator that has received little empirical follow-up is the *format of a quantifier*: whether the amount presented with the attribute is in numerical (e.g., '75% lean') or verbal (e.g., 'high % lean') format. Previous work suggests that using a verbal quantifier could magnify the framing effect size compared to numerical quantifiers (Welkenhuysen et al., 2001). However, this work has yet to be replicated on a larger scale. The goal of this paper was to test how quantifier format would moderate the framing effect based on previous empirical evidence and the predicted theoretical properties of verbal quantifiers.

6.2.1 The role of quantifier format in the framing effect

A framing scenario is typically constructed using numerical quantifiers, with which it is easy to create mathematical complements. However, some studies showed that the framing effect also occurred with verbal quantifiers (Welkenhuysen et al., 2001; see also Reyna & Brainerd, 1991, for an example in risky choice framing). Past work showed that verbal quantifiers produced a greater framing effect than their numerical equivalent (Welkenhuysen et al., 2001). In this case, Welkenhuysen et al. (2001) found that participants wanted to take a diagnostic test for cystic fibrosis more with the verbal negative frame (chance of a baby

[1] Attribute framing is one of a larger class of framing effects, involving risky choice framing and goal framing (for a review of framing typology, see Levin et al., 1998). In this paper, we refer specifically to the framing effect in *attribute framing* only.

with cystic fibrosis) than the verbal positive frame (chance of a baby without the disease). The effect was not found with the numerical frames.

While this empirical evidence suggests that verbal quantifiers can magnify the framing effect, two issues constrain such a conclusion. First, the previous study used verbal translations generated by a pilot sample for two numerical attribute frames (positive frame: 75% or high chance of a baby without cystic fibrosis; negative frame: 25% or moderate chance of a baby with cystic fibrosis). Individuals vary in their interpretations of verbal quantifiers (e.g., Berry et al., 2002; Budescu et al., 2012), making it difficult to ensure that a verbal quantifier is uniformly interpreted between participants and across frames.

Second, the previous study used the context of genetic counselling, where attitudes may have already been predisposed towards prenatal testing (Decruyenaere et al., 1992; Janssens et al., 2016). It is thus uncertain if the format effect would translate to a context where attitudes are more neutral —for instance, the traditional attribute framing manipulation (Levin & Gaeth, 1988). Therefore, the experiments in this paper sought to test the moderating role of quantifier format in a traditional attribute framing context, using a design that controlled for individual variation in verbal-numerical quantifier translations.

6.2.2 Why verbal quantifiers should magnify the framing effect

From a theoretical perspective, two proposed properties of verbal quantifiers predict that they should produce a larger framing effect than numerical quantifiers. First, verbal quantifiers could increase intuitive biases compared to numerical ones (Windschitl & Wells, 1996). One explanation posited for the attribute framing effect is that an intuitive response to the positive affect for the 'lean' frame creates a positive bias towards this frame compared to the 'fat' frame (Levin, 1987). The link between analytical (i.e., non-intuitive) processing and reduced decision biases has been demonstrated in risky choice framing scenarios (Keysar et al., 2012; Thomas & Millar, 2012). It is reasonable that greater intuitive responding would also result in greater attribute framing bias, especially as the associative processes that may account for the attribute framing effect are linked to intuitive judgements (Morewedge & Kahneman, 2010). Therefore, if verbal quantifiers are indeed more intuitive, they should magnifying the framing effect compared to numerical quantifiers by increasing the automatic affective response to the frame: a high % of lean meat is more positive than a 75 % of lean meat because people intuitively respond more to the verbal positive frame (and vice versa for fat).

Second, verbal quantifiers possess an inherent quality additional to the amounts they express: verbal quantifiers also direct a reader's focus to an attribute (Sanford & Moxey, 2003). The focusing property of verbal quantifiers is posited to be greater than that of numerical quantifiers (Teigen & Brun, 1995, 2000). Sanford et al. (2002) argued that framed numerical quantifiers also result in changes of focus, as they found that for some numerical quantifiers, participants took longer to read sentences that included a conclusion incompatible with the focus of the frame. Specifically, their participants took longer to read that a 5% fat yoghurt was unhealthy than a 5% fat yoghurt was healthy (Sanford et al., 2002). However, this effect was not found with a 25% fat (75% lean) yoghurt. Furthermore, the difference in reading times could be due to the frame's, rather than the numerical quantifier's focus. In addition, the study did not test a verbally quantified frame, which would likely have a stronger focus. More recently, Moxey (2017) showed that readers detected changes in attributes more between statements with verbal than numerical quantifiers, which suggests that verbal quantifiers do put a greater focus on the attribute. As such, if beef described with a 'high % of lean meat' increases focus on the leanness of the meat compared to beef described with '75% of lean meat', one might encode the verbally-quantified information about the meat's leanness more deeply (Moxey, 2017). This would result in a more positive judgement of the high % lean beef than the 75% lean (and vice versa for fat), and thus a magnified framing effect for verbal quantifiers. We were therefore interested in whether a stronger focus could explain a magnified verbal framing effect.

6.2.3 Present research

The three experiments reported herein were designed to systematically address whether the framing effect increases when a verbal, rather than numerical, quantifier is used. We also tested whether this might be explained by focusing properties of verbal quantifiers (Sanford & Moxey, 2003), which we hypothesised would be greater than those of numerical quantifiers (Moxey, 2017; Teigen & Brun, 1995, 2000). We used the attribute framing context of fat vs. lean meat (Levin, 1987; Levin & Gaeth, 1988), which has been replicated in many independent studies (Donovan & Jalleh, 1999; Kim et al., 2014; Kreiner & Gamliel, 2017; Seta et al., 2010). The framing effect has been extended to different quantity pairs with a range of contrasting complementary values (e.g., 25% vs. 75%, 20% vs. 80%) and the findings are inconsistent as to whether the different mathematical complements cause a larger or smaller effect (studies reporting different effect sizes across pairs with different complements: Janiszewski et al., 2003; Kim et al., 2014; Sanford et al., 2002; studies reporting no differences: Jin et al., 2017; Olsen, 2015). We therefore tested a range of quantity pairs to ascertain if the predicted larger verbal framing effect would be robust across different complements.

In line with most recent scientific guidelines, all our methods and hypotheses were pre-registered prior to conducting the experiments. The pre-registrations, materials, and data for the experiments are available on the Open Science Framework.

6.3 Experiment 1

6.3.1 Method

Participants. The experiment was powered to detect an interaction with an expected effect size of f = .10 (α= .05, 1-β= .80, minimum required sample size was 280 participants). Participants were sourced from a survey panel company (N = 363; offered online vouchers for participation) and from a university undergraduate pool (N = 181; rewarded with course credit). After excluding unfinished and careless responses according to *a priori* defined criteria (either finishing in less than one-third the median completion time or failing to disagree with the attention check question, '*I have never brushed my teeth.*'), the sample had 335 participants (194 from the survey panel, 161 undergraduates)[2]. Participants were 59% female, 80% White, with an age range of 18-76 years (M = 37.76, SD = 17.30). They had on average a healthy BMI (M = 24.99, SD = 5.78) and slightly positive attitudes towards healthy eating (M = 4.89 out of 7, SD = 0.98). Fifty-three percent reported frequent use of nutrition labels.

Design. The experiment used a 2 (frame: positive [lean] vs. negative [fat] × 4 (quantity pair: four complements, see Table 6.1) × 2 (format: verbal vs. numerical) mixed design, with frame and format manipulated between-subjects and quantity magnitude within-subjects. Order of presentation was randomised.

Materials and procedure. After providing informed consent, participants performed a translation task, where they provided verbal equivalents of the numerical quantities in the experiment. The purpose of this task was to create equivalent verbal and numerical frames for comparison across conditions. Participants selected for four quantity magnitudes the most appropriate verbal quantifier from a randomised drop-down list of 13 (see Appendix E for the full list of verbal quantifiers). For example, they were told '*the beef contains 25% of fat meat*', and they picked a word to complete the sentence: '*There is a(n) __________ percentage of fat meat in the beef.*' Participants also provided two filler translations of other food quantities (e.g. 'low % calories') that served as distractions. Table 6.1 shows the most common translation of numerical quantities into verbal ones. We subsequently presented participants with their selected verbal quantifier in the verbal condition of the framing task.

After performing the translation task, participants completed a distraction task that required them to complete a sentence describing computer battery life or jeans shrinkage similar to the ones used in Teigen et al. (2014).

After the distraction task, participants performed the following framing task for four different quantity pairs (presented in randomised order). Participants judged the healthiness of meat in the following vignette:

[2] The framing effect was not significantly different between samples, thus all results were analysed with both samples combined.

Table 6.1. Most common translations of numerical quantities into verbal quantifiers.

Numerical quantity	**Most common verbal translation range**	**Percentage of participants**		
		Exp 1 (*N* = 335)	**Exp 2 (*N* = 442)**	**Exp 3 (*N* = 440)**
Positive frame				
95% lean	Very high	75	63	84
75% lean	High	53	40	52
50% lean	Medium	76		77
25% lean	Low	62		67
Negative frame				
5% fat	Low	58	58	51
25% fat	Medium	40	37	48
50% fat	Very high	49		41
75% fat	Very high	77		68

You are given the following information about a 250g beef fillet:
The beef contains a [quantity] % of [attribute] meat.

Participants made healthiness judgements on a Likert scale (1: very unhealthy, 7: very healthy). They also indicated how much they would be willing to pay (WTP) for the meat (in pounds sterling). We excluded from analysis WTPs that were more than five standard deviations above the mean. Participants also completed this task for two filler items (cereal bars with different energy values).

Finally, participants reported their attitudes towards healthy eating (Steptoe et al., 1995), how frequently they used nutrition labels, and socio-demographic information, including weight and height.

6.3.2 Results

Traditional framing effect. We observed the traditional framing effect over all conditions: participants rated the % lean meat as healthier and were willing to pay more for it than % fat meat. Figure 6.1 illustrates the mean

Table 6.2. Magnitude of framing effect (positive frame minus negative frame) for healthiness judgements across quantity pairs in Experiments 1-3.

Quantity pair	**Framing effect magnitude (Mean judgement difference between lean and fat frame)**					
	Verbal		**Numerical**		**Overall**	
	M	***95% CI***	***M***	***95% CI***	***M***	***95% CI***
Experiment 1 (7-point scale; N = 355)						
5% fat	0.53*	[0.10, 0.96]	-0.01	[-0.44, 0.42]	0.26	[-0.05, 0.56]
25% fat	2.10***	[1.67, 2.53]	1.11***	[0.68, 1.54]	1.60***	[1.30, 1.91]
50% fat	1.40***	[1.02, 1.78]	1.49***	[1.11, 1.87]	1.45***	[1.18, 1.72]
75% fat	0.83***	[0.39, 1.28]	0.89***	[0.45, 1.34]	0.86***	[0.55, 1.18]
Experiment 2 (100-point scale; N = 442)						
5% fat	3.65	[-1.37, 8.68]	4.84	[-9.84, 0.16]	4.25*	[0.71, 7.79]
25% fat	26.95***	[21.03, 32.87]	34.65***	[28.76, 40.54]	30.80***	[26.62, 34.98]
Experiment 3 (11-point scale; N = 440)						
5% fat	0.30	[-0.21, 0.82]	0.92**	[0.40, 1.43]	0.61**	[0.24, 0.97]
25% fat	2.40***	[1.84, 2.96]	3.04***	[2.47, 3.60]	2.72***	[2.32, 3.12]

Note. Larger scores indicate a larger framing effect. Results of significance testing are given for p < .05*, p < .01**, and p < .001***. The differences in framing effect between verbal and numerical quantifiers are illustrated in Figures 6.1 and 6.2

distributions of healthiness judgements and willingness-to-pay values for each of the four contrasting quantities. Table 6.2 shows the mean difference in healthiness judgements (positive –negative) for the different quantity pairs: scores further from zero indicate stronger framing effects. As healthiness and willingness-to-pay were significantly correlated, we conducted a mixed MANOVA on healthiness judgements and willingness-to-pay, using frame and format as between-subjects factors and quantity pair as a within-subjects factor. The framing effect was only significant for healthiness judgements, and not for willingness-to-pay, $F(1, 331) = 99.40$, $p < .001$, $\eta^2_P = .23$; $F(1, 331) = .06$, $p = .815$, $\eta^2_P < .001$, respectively.

Quantifier format moderated the framing effect for only one out of four quantity pairs. Format did not have a main effect on healthiness ratings or WTP, $F(1, 331) = .01$, $p = .917$, $\eta^2_P < .001$; $F(1, 331) = .25$, $p = .618$,

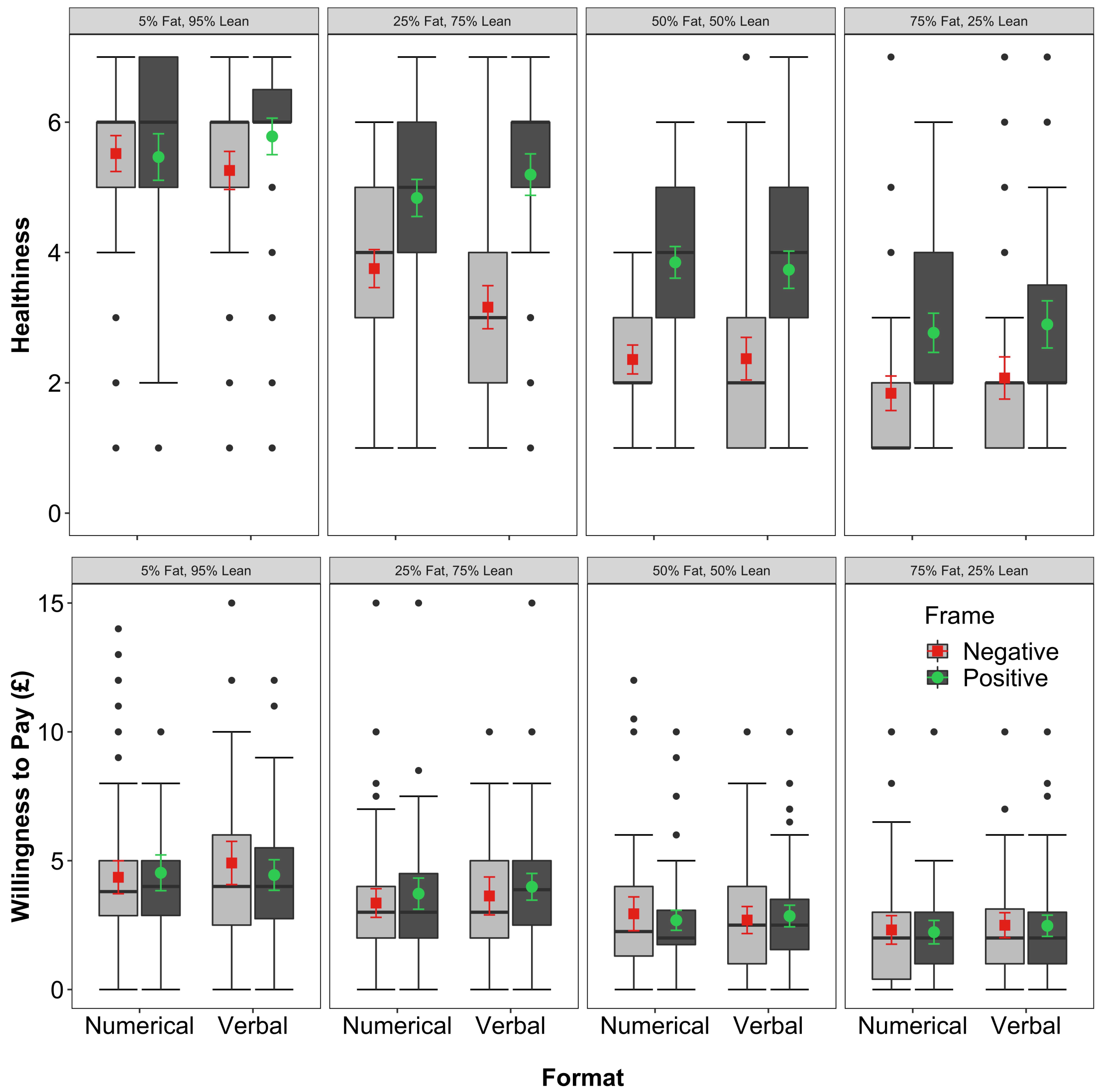

Figure 6.1. Means and distributions of healthiness judgements and willingness-to-pay values for verbal and numerical quantifiers across four quantity pairs in Experiment 1.

Note. The point plots (red and green) give the means and 95% confidence intervals of participants' responses, and the box-and-whisker plots show the overall distributions of the responses. Framing effects (difference between positive and negative frame) were only significant for healthiness judgements.

η^2_P = .001, respectively. However, the framing effect for verbal and numerical quantifiers differed across the four quantity pairs (see Table 6.2 and Figure 6.1), $F(3, 975) = 4.05$, p = .007, η^2_P = .01. We followed up on the significant three-way interaction to test whether format moderated the framing effect for each quantity pair. The 75% lean vs. 25% fat pair produced a significantly larger framing effect in the verbal than the numerical format, $F(1, 331) = 9.49$, $p = .002$, $\eta^2_P = .028$. However, none of the other three quantity pairs (5% fat, 50% fat, and 75% fat) showed a significant moderating effect of format, $F(1, 331) = 3.54$, $p = .061$, $\eta^2_P = .01$ (5% fat); $F(1, 331) = 0.21$, $p = .647$, $\eta^2_P = .001$ (50% fat); $F(1, 331) = 0.11$, $p = .741$, $\eta^2_P = < .001$ (75% fat).

Variations in framing effect across quantity pairs. As illustrated in Figure 6.1, the size of the framing effect varied across the four quantity pairs, with the largest framing effect observed for the 25% fat pair, and the smallest with the 5% fat pair (see Table 6.2). This was also quantified by significant two-way interactions between frame and quantity pair for healthiness and willingness-to-pay, $F(3, 975) = 22.91$, $p < .001$, $\eta^2_P = .07$ and $F(3, 975) = 3.48$, $p = .015$, $\eta^2_P = .01$, respectively.

6.3.3 Discussion

Experiment 1 only found limited support for our hypothesis that quantifier format would moderate the framing effect. Only the 25% fat quantity pair had a significantly larger verbal than numerical framing effect on healthiness judgements. We also found that the framing effect size was not consistent across the different quantity pairs. The largest framing effect was found with the 25% fat pair, and the smallest (no significant effect) with the 5% fat pair —though this was primarily due to the lack of framing effect with the numerical quantifier in this condition. One potential explanation for the reduced framing effect with this quantity pair (smaller in size for low %, and non-existent for 5%) is that the healthiness judgements were too close to ceiling, as this would be the healthiest beef in the set. Another explanation for the smaller framing effect is that certain quantifiers produce a focus that directs a reader away from the attribute rather than to it (Teigen & Brun, 2003): a 'low' or 5% fat beef may indicate an absence of fat meat, as opposed to a presence of fat meat (as might be the case with 'moderate' or 25% fat beef). The small amount of fat meat (directing away from fat) would point in the same direction as the large amount of lean meat, reducing the framing effect. This would suggest that directional focus magnifies the framing effect if foci for both frames are in the same direction, but reduces the effect if the foci are in opposite directions. It is surprising that this would be even more the case for the numerical quantifier, but as our test of the format and frame interaction was not significant, we cannot conclude that the numerical quantifier did have a significantly smaller effect than the verbal. To address these possibilities, we sought to compare moderated framing effects between 5% fat meat (the smallest effect) and 25% fat meat (the largest effect) using a more sensitive scale, and investigate whether the focus of the quantifiers explained a moderated framing effect.

6.4 Experiment 2

The goal of Experiment 2 was to test the effect of quantifier format on attribute framing, and to explain the effect and its variation. Two posited properties of verbal quantifiers suggested that they should produce a larger framing effect than numerical quantifiers. First, verbal quantifiers are more intuitive than numerical ones (Windschitl & Wells, 1996), and second, verbal quantifiers having more focusing properties than numerical ones (Teigen & Brun, 1995, 2000). In addition, some verbal quantifiers may produce a different focus than their corresponding numerical quantifier (Teigen & Brun, 2003). For instance, 'low % fat' may indicate the absence of fat, whereas '5% fat' may still focus on what fat there is (Teigen & Brun, 2003). Focusing properties may contribute to the framing effect (Sanford et al., 2002), therefore we assessed the focus of the verbal and numerical quantifiers in the frame, and whether this focus was a contributing factor to the framing effect in general, and to an expected larger framing effect with verbal quantifiers.

To simplify our analysis, we used two quantity pairs that produced the greatest difference in framing effect size in Experiment 1: the 5% fat and 25% fat pairs. Choosing pairs with differing framing effect sizes allowed us to test whether the quantifiers in each frame had different focusing properties, and whether this contributed to the difference in effect size. In addition, we sought to rule out a methodological artefact for differences in framing effect size. We accounted for the possibility that the 7-point Likert scale in Experiment 1 might lack sensitivity and result in a ceiling effect for the 5% fat quantity pair (Voutilainen et al., 2016) by changing the response scale to allow for finer-grained measurements.

6.4.1 Method

Participants. The experiment was powered to capture the interaction effect obtained in Experiment 1 (f = .10, α= .05, 1-β= .80, minimum required sample size was 433). Four hundred and forty-two participants (72% female; 90% White; age range 18-80 years, M = 35.98, SD = 10.98) were offered £0.60 to complete the 5-minute experiment on Prolific Academic[3]. We used the same exclusion criteria as in Experiment 1. Participants had on average an overweight BMI (M = 27.79, SD = 8.99) and positive attitudes towards healthy eating (M = 5.10, SD = 0.84); seventy-three percent reported frequent use of nutrition labels.

Design. The design was the same as Experiment 1 (format and frame manipulated between-subjects; quantity pairs within-subjects), however we only used the 5% and 25% fat pairs.

Materials and procedure. After providing informed consent, participants performed the translation task for the numerical quantifiers as in Experiment 1 (see Table 6.1). They then rated the healthiness of meat described in the same Experiment 1 vignette for each quantity pair (shown in randomised order to each participant) on a sliding scale that increased from 0-100 in invisible increments of 1. Seven descriptors were spaced over the scale (from very unhealthy to very healthy). After this, participants were presented again with the vignettes and given the following sentence completion task:

Pick the option that makes the most sense to complete the sentence:
The beef has [quantifier] % of [attribute] meat because ...
A the cow was grain-fed and developed a lot of fatty tissue [fat focus]

[3]This payment amounts to a £7.20 per hour wage, which is above the minimum wage recommendation for survey panel studies.

B *the cow was grass-fed and developed a lot of lean muscle* [lean focus]

The two options were presented in a random order. One option always focused on the presence of the attribute in the vignette (e.g., option A in the example for fat meat), while the other focused on its absence (e.g., option B for fat meat; vice versa for lean meat). At the end of the experiment, participants completed the same demographic survey as in Experiment 1.

6.4.2 Results

Does quantifier focus explain the framing effect?

We ran pre-registered mediation analyses to assess whether the effect of frame on healthiness judgement was mediated by quantifier focus (1000 simulations using the R package 'mediation'; Tingley et al., 2014). This allowed us to estimate and test the average causal mediation effect and average direct effect as moderated by quantifier format for each of the two quantity pairs. The middle columns of Table 6.3 report the mediation analyses for each quantity pair. The results indicated that framing had a direct effect on healthiness judgement for both the 5% fat and 25% fat pairs. This traditional framing effect is depicted in the top panel of Figure 6.2, and showed that participants judged the lean meat as healthier than the equivalent quantity presented in terms of fat content. The effect of framing was significant overall for both the 5% fat and 25% fat pairs, b = 4.22, p = .022, 95% CI [0.66, 7.61] (5% fat); b = 30.67, p < .001, 95% CI [26.42, 34.78] (25% fat).

Role of format in the mediation of the framing effect by focus. The moderated mediation analyses showed that the framing effect was not significantly different between verbal and numerical format, whether it was the direct effect of frame on healthiness or mediated by focus. Therefore, contrary to our expectations, quantifier format did not result in variations in the framing effect.

Role of quantifier focus as a mediator of the framing effect. We expected participants to focus more on the attribute in the frame rather than its complement, as indicated by causal completions that explained the frame in terms

Table 6.3. Magnitude of framing effect (positive frame minus negative frame) for healthiness judgements across quantity pairs in Experiments 1-3.

	Experiment 2		**Experiment 3**	
Framing effects	**5 (low) % fat vs. 95 (very high) % lean**	**25 (moderate) % fat vs. 75 (high) % lean**	**5 (low) % fat vs. 95 (very high) % lean**	**25 (moderate) % fat vs. 75 (high) % lean**
Total framing effect	4.22 [0.66, 7.61]	30.67 [26.42, 34.78]	0.60 [-0.11, 0.30]	2.71 [2.32, 3.11]
Effect on focus to lean frame (direct)	-0.82 [-2.15, 0.33]	2.09 [1.49, 2.72]	0.47 [-0.81, 1.86]	2.61 [1.94, 3.35]
Effect on healthiness (direct)	4.82 [1.50, 8.02]	22.87 [18.43, 27.31]	0.55 [0.18, 0.92]	1.63 [1.23, 2.08]
Indirect effect (causal mediation)	-0.60 [-2.17, 0.97]	7.80 [5.05, 10.66]	0.05 [-0.05, 0.16]	1.08 [0.79, 1.39]
Tests of moderated mediation				
Direct effect	-0.03 [-6.55, 6.54]	-7.27 [-16.00, 1.84]	-0.58 [-1.31, 0.15]	-0.59 [-1.46, 0.15]
Indirect effect	-1.23 [-4.26, 1.61]	-0.49 [-5.22, 4.00]	-0.02 [-0.20, 0.15]	-0.05 [-0.52, 0.41]

Note. 95% confidence intervals (indicated in square brackets) were generated using 1,000 bootstrap samples. Focus on the lean frame was tested as a mediator of the frame-healthiness relationship.

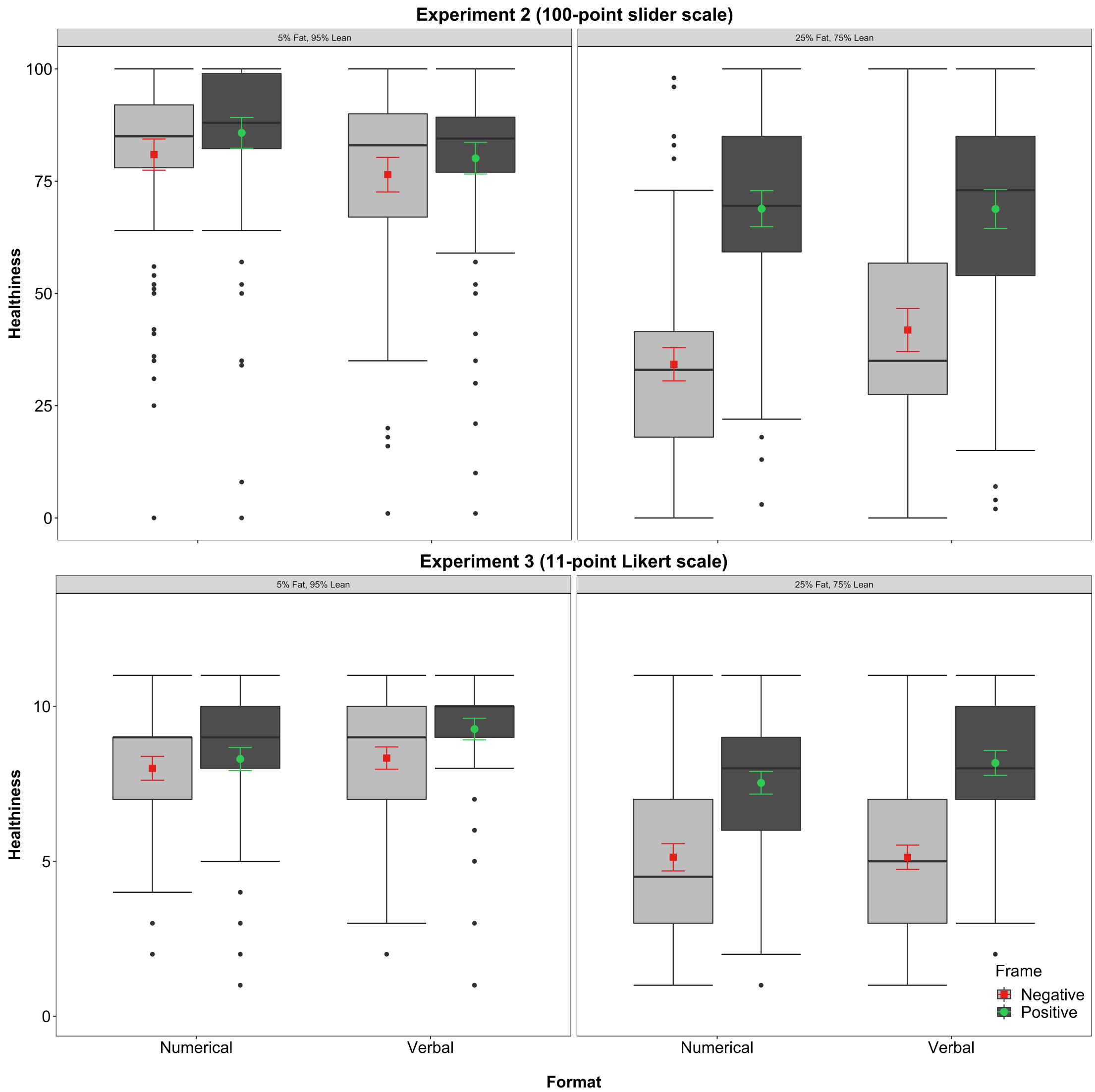

Figure 6.2. Means and distributions of healthiness judgements for verbal and numerical quantifiers in the 5% and 25% fat pairs in Experiments 2 (100-point scale) and 3 (11-point scale).

Note. The point plots (red and green) give the means and 95% confidence intervals of participants' responses, and the box-and-whisker plots show the overall distributions of the responses. Framing effects (difference between positive and negative frame) were significant for the 25% fat but not 5% fat pair.

of the presented attribute. As illustrated in the top panel of Figure 6.3, we found that participants did select more sentence completions with a lean focus when the beef was described as 75 (or high) % lean, but more sentence completions with a fat focus when the product was described as 25 (or moderate) % fat. However, in the other quantity pair, where the beef was described as 95 (or very high) % lean or 5 (or low) % fat, participants selected more sentence completions with a lean focus regardless of the frame. The mediation analysis showed that a greater focus on the lean attribute mediated the effect of framing on healthiness judgement for the 25% fat pair, but not the 5% fat pair, b = 7.80, $p < .001$(25% fat), 95% CI [5.05, 10.66]; b = -0.60, p = .406, 95% CI [-2.17, 0.97].

6.4.3 Discussion

Experiment 2 gave more evidence of the robustness of the framing effect, but also that using a verbal quantifier did not magnify the effect of framing on healthiness judgements. We expected that the choice of attribute frame would lead a reader to focus on reasons justifying the attribute cited (e.g., 'lean' in a lean frame), and that this would explain the framing effect. Our evidence supported this mediating role of focus for the 25% fat pair, where both the numerical and verbal quantifiers had similar focus on the fat or lean frame respectively. However, we did not find the same effect with the 5% fat pair: in both numerical and verbal condition, participants focused on explaining how lean the product was, no matter whether it was described with a fat or lean frame.

The fact that verbal quantifiers did not lead to larger framing effects in Experiment 2 could indicate that verbal quantifiers do not magnify framing effects compared to numerical ones. However, two factors in the experiment constrain this conclusion. First, we used a verbal to numerical translation task at the onset of the study that may have primed people to think about verbal quantifiers in a numerical way. This could have rendered verbal statements more similar to numerical ones. Second, the 100-point response scale, which successfully increased sensitivity to detect smaller judgement differences for the 5% fat pair, could have inadvertently caused an anchoring of judgements in the numerical condition to the corresponding scale points (e.g., 25% fat is 25/100 healthy), thereby widening the response range between the numerical frames. In our next experiment, we sought to address these concerns.

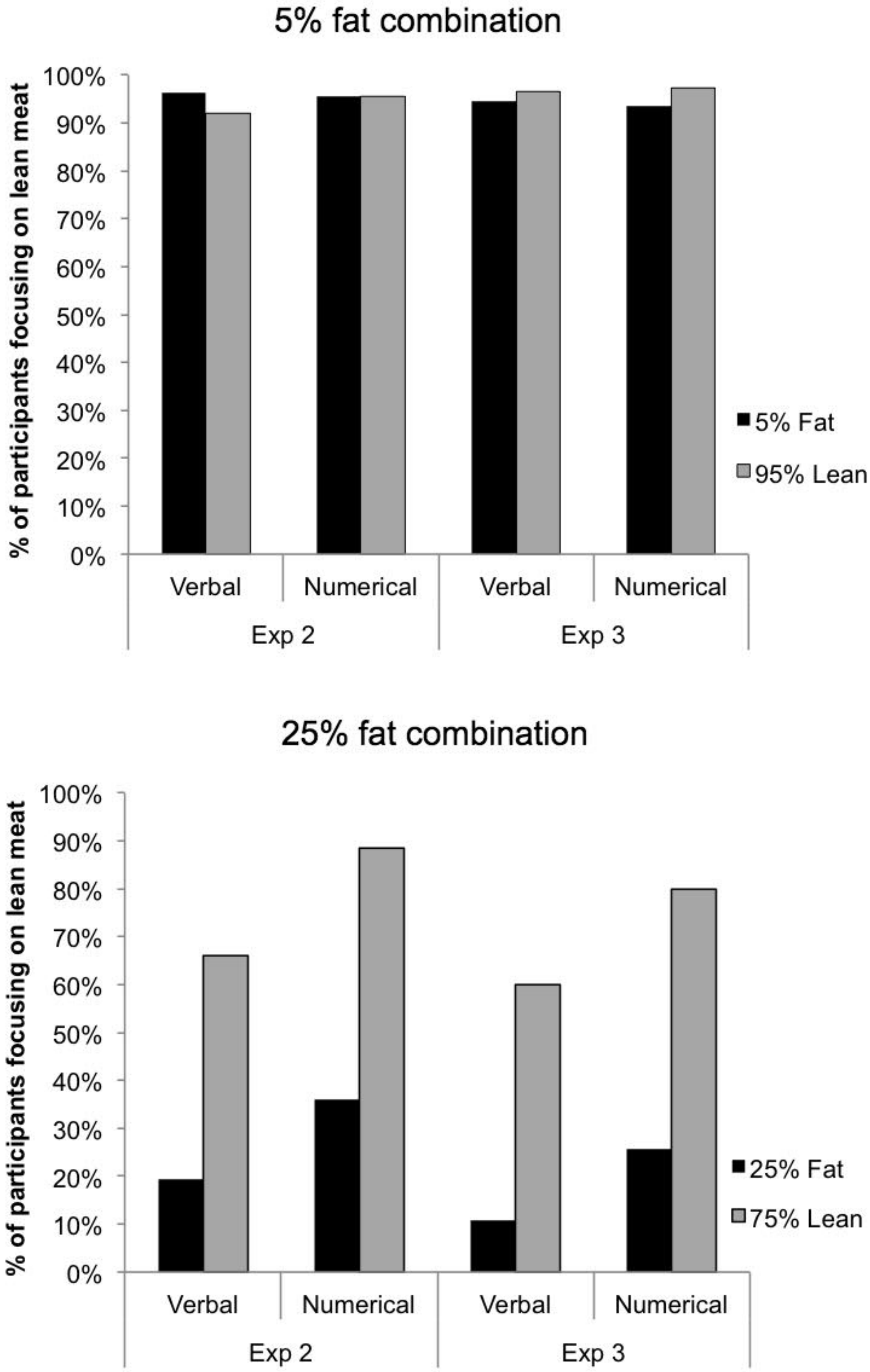

Figure 6.3. Percentage of participants who inferred reasons congruent with the presence of lean meat, by format, frame, and quantity combination.

Note. The corresponding percentage would be the percentage of participants who inferred reasons congruent with the presence of fat meat.

6.5 Experiment 3

Experiment 3 aimed to replicated Experiment 2 while overcoming two methodological limitations. First, we re-introduced a distractor task between the translation task and the actual framing evaluation task to reduce the likelihood that people were still thinking about their translations. Second, we adjusted the response scale to an 11-point Likert scale to reduce anchoring of responses to scale values. We tested again whether verbal quantifiers would magnify the framing effect, with focus on the lean attribute as a mediator. Based on the results from Experiments 1 and 2, we also predicted that the framing effect size would be larger for the 25% fat than the 5% fat pair.

6.5.1 Method

Participants. The experiment was powered to capture the previous interaction effects obtained (f = .10, α= .05, 1-β= .80, minimum required sample size was 433 participants). Four hundred and forty participants (66% female; 89% White; age range 18-74 years, M = 33.90, SD = 11.59) were offered £1 to complete the 10-minute experiment on Prolific Academic. The exclusion criteria were identical to Experiments 1 and 2. Participants had on average a slightly unhealthy BMI (M = 26.86, SD = 7.82) and positive attitudes towards healthy eating (M = 5.01, SD = 0.94). Seventy-five percent reported frequent use of nutrition labels.

Design. The design was the same as Experiment 2, with format and frame manipulated between-subjects, and quantity pairs manipulated within-subjects.

Materials and procedure. After providing informed consent, participants completed translations for the numerical quantifiers, including six filler translations (50% fat, 75% fat, and four verbal-numerical translations for low% and high % risks). Participants then completed a distractor task where they described a graph about medical treatment outcomes. Subsequently, participants performed the healthiness judgement task for the two quantity pairs (5% fat and 25% fat) in randomised order. Responses were made on an 11-point Likert scale

(1: extremely unhealthy, 11: extremely healthy) so as to maintain the greater sensitivity of the rating scale while minimising the possibility or participants anchoring responses to the numerical quantifiers given. Following their healthiness judgements, participants performed the sentence completion task used in Experiment 2. Finally, they completed the demographic survey.

6.5.2 Results

Does quantifier focus explain the framing effect? We ran preregistered mediation analyses for the effect of frame on healthiness judgement as mediated by focus on the lean attribute (vs. the fat one) and moderated by quantifier format. The right columns of Table 3 report the mediation analyses for each quantify pair in Experiment 3. Framing had a direct effect on healthiness judgement for both quantity pairs. The bottom panel of Figure 6.2 depicts the traditional framing effect, which was significant overall for the both pairs, $b = 2.71$, $p < .001$, 95% CI [2.32, 3.11] (25% fat); $b = 0.60$, $p < .001$, 95% CI [0.21, 0.97].

Role of format in the mediation of the framing effect by focus. The tests of moderated mediation showed that quantifier format did not significantly affect the framing effect size. This provided additional evidence to Experiment 2 that contrary to expectations, the verbal quantifier did not magnify the framing effect.

Role of quantifier focus as a mediator of the framing effect. We replicated the effect of frame on sentence completions from Experiment 2. As illustrated in the bottom panel of Figure 6.3, participants selected more sentence completions with a lean focus for 75 (high) % lean mean, but more sentence completions with a fat focus for 25 (moderate) % fat meat. However, participants consistently selected sentence completions with a lean focus for both frames in the 5% fat pair. A greater focus on the lean attribute mediated the framing effect on healthiness judgement for the 25% fat pair, but not the 5% fat pair, $b = 1.08$, $p < .001$, 95% CI [0.79, 1.39] (25% fat); $b = 0.05$, $p = 0.280$, 95% CI [-0.05, 0.16] (5% fat).

Does framing effect size vary across quantity pairs? We tested whether the framing effect would be larger for the 25% fat pair than the 5% fat one in a pre-registered ANOVA with frame and quantity pair as factors. As predicted, the 25% fat pair produced a greater framing effect, $F(1, 436) = 122.71$, $p < .001$, $\eta^2_P = .22$.

6.5.3 Discussion

Experiment 3 showed a similar pattern to Experiment 2. We observed a smaller framing effect for the 5% fat pair than the 25% fat one, with no significant evidence that quantifier format moderated this effect. In addition, the frame produced a focus on the cited attribute (fat or lean) in the 25% fat pair (as evidenced by sentence completions justifying those attributes). The 5% fat pair, on the other hand, produced a consistent focus on only the lean attribute, even with a fat frame. The focus on a lean frame partially mediated the relationship between frame and healthiness judgement for the 25% fat pair, suggesting that 25 (moderate) % and 75 (lean) % direct focus to the cited attribute and hence contributed to the framing effect. However, the mediation was not observed for the 5% fat pair, likely because the quantifiers in this pair had opposing foci: 95 (very high) % directed focus to the lean frame, and 5 (low) % fat also directed focus to the lean frame —away from the fat frame. This was the case for both verbal and numerical quantifiers, contradicting the assumption that verbal quantifiers possess more focusing properties than numerical ones (Teigen & Brun, 1995, 2000).

6.6 Framing Effect and the Moderating Role of Format: Data Synthesis Across Experiments

To quantify the robustness of the framing effect and the role of format as its moderator, we meta-analysed the moderated framing effect for the 5% fat and 25% fat pairs across the three experiments reported here. Meta-analytical methods provide more precision in the estimation and minimise the chance of obtaining null effects due to lack of statistical power (Cumming, 2013). We computed the internal meta-analysis using random effect models (a restricted maximum likelihood estimator) with the R package ‘metafor’ (Viechtbauer, 2010).

The overall framing effect was significant (see Figure 6.4), $b = 0.75$, $p < .001$, 95% CI [0.42, 1.07]. Format was not a significant moderator across studies for either quantity magnitude, 5% fat: $b = 0.01$, $p = .934$, 95% CI [-0.31, 0.33]; 25% fat: $b = 0.12$, $p = .643$, 95% CI [-0.40, 0.65]. Thus, we did not find evidence to support our hypothesis that verbal quantifiers would exhibit a larger attribute framing effect than numerical quantifiers.

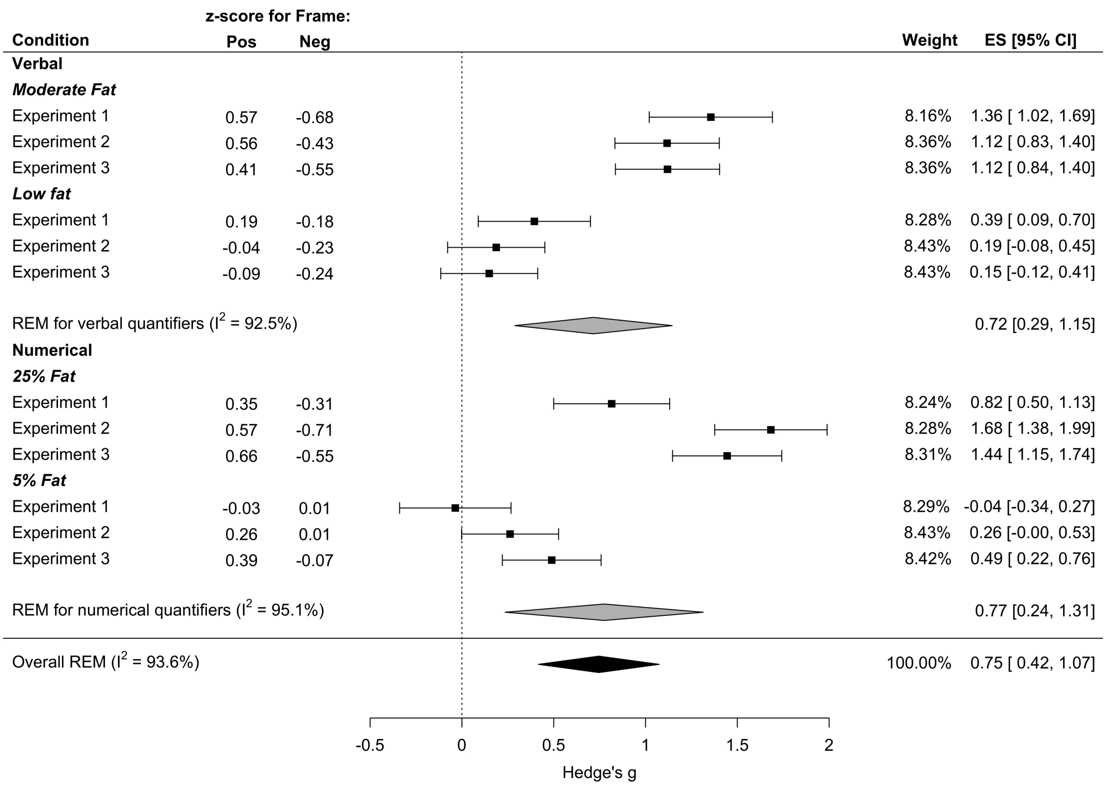

Figure 6.4. Forest plot of the framing effect sizes for verbal and numerical conditions across three experiments showing similar framing magnitude across formats.

Note. The grey diamonds show the random effects model for verbal and numerical quantifiers, and the black diamond shows the overall effect size across all formats and quantity combinations.

6.7 General Discussion

In three experiments, we investigated how the format of a quantifier moderates the attribute framing effect across different quantity combinations, and whether the directional focus of a quantifier could explain the effects. Across the three experiments, we replicated the traditional framing effect, showing that meat described in terms of its lean content was judged healthier than meat described with an equivalent fat content, but we did not find evidence that verbal quantifiers magnify the framing effect.

6.7.1 No effects of quantifier format: Implications for previous empirical findings

Contrary to our predictions and previous empirical findings (Welkenhuysen et al., 2001), verbal quantifiers only magnified the framing effect for 25% fat in Experiment 1, and this result was not replicated in Experiments 2 and 3. Our investigation used three well-powered experiments and the original attribute framing design in the literature (Levin, 1987) to test for the format moderation effect. In addition, we controlled for individual variation in how people translate between numerical and verbal quantifiers by using each participant's own translations in the framing scenarios. A meta-analysis across the three experiments found that the overall moderating effect of format was not significant, and it was close to zero. Previous findings of a larger framing effect with verbal quantifiers (Welkenhuysen et al., 2001) may have reflected translation differences between a pre-test and experimental sample, or the specificity of the context in which it was tested. Further investigations may consider whether there are elements of the context that may still produce the moderating effect of format in attribute framing, when translation differences are accounted for.

6.7.2 Implications about quantifier properties and framing effects.

We predicted that verbal quantifiers would magnify the framing effect compared to numerical ones based on two posited properties of verbal quantifiers. First, that verbal quantifiers are processed more intuitively and therefore lead to more judgement biases (Windschitl & Wells, 1996). Second, that verbal quantifiers have greater focusing properties that would lead to stronger focus on the attribute and greater pragmatic signals that this attribute is important to the judgement (Teigen & Brun, 1995, 2000). Finding that the quantifier format did not moderate the framing effect could either mean poor evidence that these properties differ between verbal and numerical quantifiers, or poor evidence for the assumption that these properties produce the framing effect. To address this issue, we examine two findings of our data: the variations in framing effect size between different quantity pairs, and the mediation of the framing effect by quantifier focus.

Classic framing effect is robust but varies in effect size. The classic framing effect is considered a judgement bias since the two frames are logically equivalent, so by rational reasoning, one should not prefer one to the other (Tversky & Kahneman, 1986). We replicated the classic effect, with similar effects between verbal and numerical quantifiers. This suggests that the quantifier format did not significantly affect participants' processing of the overall information, since they produced the same biases. In addition, we found variations in the framing effect size among quantity pairs: the 25% fat pair consistently produced the largest framing effect. This supports work that found effect size variations across frames with different quantity pairs (Janiszewski et al., 2003; Kim et al., 2014; Sanford et al., 2002), and contradicts work that did not (Jin et al., 2017; Olsen, 2015). Assuming that the framing effect is an intuitive bias driven by an initial affect response to the positive or negative frame (Levin, 1987), one could expect a similar framing effect size irrespective of the exact quantity pair because the association created by 'fat' (or 'lean') is present in every pair. The differences in effect sizes between quantity pairs like 25% vs. 5% fat (75% vs. 95% lean) suggests some additional processing to reach a judgement. If we consider intuitive processing to produce fast, automatic responses (De Neys, 2017b; Evans, 2008), additional processing about the quantity would be incompatible with intuition.

One could argue that the quantifier could automatically scale the affective reaction to the frame: for example, 'low fat' might produce an instinctively positive affective response rather than a negative one. If this were the case, we would also expect to see the scaling extend similarly across frames: moderate (or 25%) fat might be more negative than low (or 5%) fat, but very high (or 95%) lean should also be more positive than high (or 75%) lean. Even if the scaling is asymmetric for positive and negative frames (e.g., people are more averse to losses than they are receptive to gains; Tversky & Kahneman, 1991), we should expect to observe this asymmetry in the quantity pairs in Experiment 1. The negative-ness of 50% fat and 75% fat should loom larger than the positive-ness of 50% and 25% lean, which should have produced a larger framing effect than for the 25% fat (or 75% lean) pair. However, this was not the case in Experiment 1. It is also worth noting that when translating numerical to verbal quantifiers, both 50% and 75% fat were most commonly described as 'very high', but different verbal quantifiers were used for all four quantities of lean meat. Participants thus seemed to be more sensitive to gradation in the lean attribute than the fat one. A focus on the lean attribute might therefore better explain the framing effect.

Focus on an attribute partially mediates the framing effect. An alternative perspective to the framing-as-bias account of framing effects is that frames are a practical, and therefore rational, source of information for a judgement (Sher & McKenzie, 2008). For example, one might infer that a lean frame was chosen because the speaker *wanted* the listener to focus on that attribute (Keren, 2007). Further, the quantifier can provide more information about what the focus should be: a quantifier can direct focus to the cited attribute (e.g., '*a few* snacks were healthy' puts focus on the healthy snacks), or away from it (e.g, '*few* snacks were healthy' puts focus on the snacks that were unhealthy; Sanford et al., 1996). Based on the sentence completion task, our participants found explanations focusing on fat to be more reasonable for 25 (moderate) % fat frames, but not for 5 (low) % fat ones. This difference in focus was observed for both verbal and numerical quantifiers, which contrasts with previous research that found numerical probabilities such as 'a 30% chance' to have more ambiguous focusing properties than their average translated verbal probability (e.g., 'unlikely', which focuses strongly on the non-occurrence of an event; Teigen & Brun, 2000). By examining the effect of a frame on focus for both the verbal and numerical quantifier, we were able to disentangle the focusing property of a frame from that of a quantifier, and the evidence for the 5 (low) % quantity pair suggests that numerical quantifiers *can* unambiguously focus away from an attribute, just like the verbal quantifier.

Our mediation analyses found that a greater focus on the lean attribute (compared to the fat one) explained the framing effect for the 25% fat quantity pair but not the 5% fat one. Taken together with the consistently larger framing effect for the 25% fat pair, the mediation results suggest that the focus is only a helpful predictor of the framing effect under certain conditions. One key feature of the 25% fat pair that distinguishes it from the others is that it is a more ambiguous complement pairing, that may be less immediately informative about its position on a scale of healthiness (e.g., people may be uncertain about what exactly is a healthy level of fat; Diekman & Malcolm, 2009). People tend to draw more from implicit information (i.e., pragmatic inferences) when they need to distinguish ambiguous targets (Grodner & Sedivy, 2011). Given a situation illustrated in Figure 6.5, one may have a vague idea of the range of fat quantities that might be considered healthy, but be uncertain whether 25% (or moderate) falls within that range (Janiszewski et al., 2003). One might then rely on the implicit focus in the quantifier and frame to infer that 25% is a larger than usual amount of fat, and thus not so healthy (Donovan & Jalleh, 1999; Sher & McKenzie, 2006). This process would be reversed in the lean condition, resulting in a conclusion that the 75 (high) % lean beef is healthier. In contrast, a 5% fat (95% lean) beef is more apparently a healthy quantity, meaning the frame and focus is less informative to the judgement. Such an explanation would also fit evidence that knowledge about typical quantity ranges (which reduces informational ambiguity) reduces the framing effect (Leong et al., 2017). Although we did not formally assess people's existing knowledge of the typical range of fat in meat, it is reasonable to assume that 5 (low) % fat is more clearly healthy than 25% fat. A salient question for future research is whether manipulating ambiguity about a quantifier could eliminate or magnify the framing effect. This would help to ascertain the conditions under which focusing properties best explain the framing effect.

6.7.3 Conclusion

The three experiments reported in this paper showed that contrary to previous empirical findings, the size of the attribute framing effect was not affected by quantifier format. Instead, the effect size varied across quantity pairs. We found evidence that the quantifier —both verbal and numerical —directed participants' focus either to or away from the attribute cited in the frame. Where participants' focus was on the cited attribute, this focus contributed to the framing effect. Taken together, these results suggest that attribute framing is more than

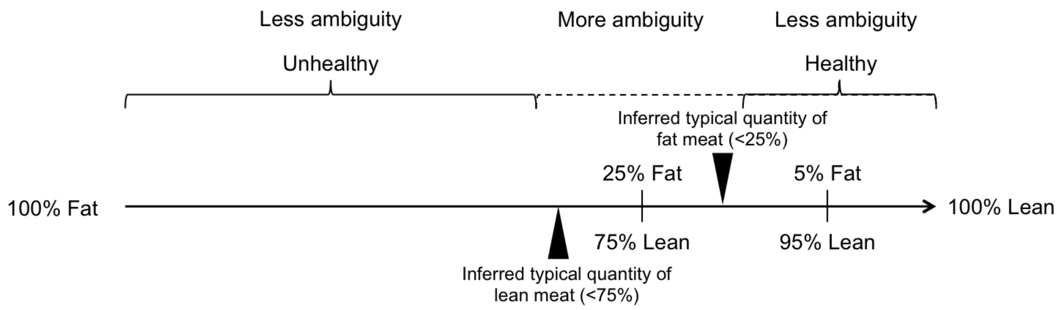

Figure 6.5. Ambiguity in subjective evaluation of framed targets.

Note. The dotted line shows a range of ambiguity for illustrative purposes. Although consumers believe that fat consumption should be less than 15% of daily calorific intake (Diekman & Malcolm, 2009), consumers' perception of the range of healthy fat percentages specific to meat is not known.

an intuitive bias: people focus their attention based on the quantifiers given, and may possibly do so more when those quantifiers are ambiguous. This supports a more rational view of framing effects.

Chapter 7: Attribute Framing with Verbal and Numerical Quantifiers

7.1 Abstract

The attribute framing effect, where people judge an item more positively with a positive attribute (e.g., ‘75% lean’) than a negative, albeit normatively equivalent attribute (e.g., ‘25% fat’), is a robust phenomenon. Existing theories and some findings from Chapter 6 predict that the framing effect is stronger when verbal quantifiers (e.g., ‘high’) are used instead of numerical ones (e.g., ‘70%’). Over four well-powered, pre-registered studies, we tested this prediction across different quantifier magnitudes in a 2 (frame: energy or calories; between-subjects) × 2 (quantifier format: verbal or numerical; between-subjects) × 2 (quantifier magnitude: small or large; within-subjects) mixed design. We also tested two mechanisms for the framing effect and its moderation, based on affective encoding and pragmatic inference accounts of the framing effect: whether the effect was mediated by the affect associated with the frame, and whether participants inferred a strong recommendation from the speaker. We found a robust framing effect, which was larger for small, but not large, verbal quantifiers. Both the affect associated with the frame and participants’ inferred strength of recommendation by the speaker partially mediated the framing effect. Both mediators were similarly strong for small quantities, but affect was the stronger mediator for large quantities. We also found that framing depended on the valence of the phrase (e.g., ‘low energy’) rather than the valence of the attribute (e.g., ‘energy’). These results suggest that both affective associations and pragmatic inference about quantifiers and frames are involved in attribute framing.

7.2 Introduction

The attribute framing effect has been one of the most widely studied psychological biases in decision-making since early research demonstrated that people prefer 75% lean meat to its mathematically equivalent 25% fat (Levin, 1987). Extensive research has documented the existence of the framing effect in many important domains (e.g., health care, Gamliel & Peer, 2010; medical risks, Peng et al., 2013; Welkenhuysen et al., 2001; business performance, Janiszewski et al., 2003; Kuvaas & Selart, 2004; sporting performance, Leong et al., 2017; enjoyment of events, Isaac & Poor, 2016; and even mate choice, Saad & Gill, 2014). The effect is robustly replicated, and some conditions appear to magnify or reduce the effect (Piñon & Gambara, 2005; Steiger & Kühberger, 2018).

In a typical attribute framing study, participants judge items described with one of two complementary phrases with a positive or negative attribute (Levin et al., 1998). For example:

The meat is 75% lean (positive frame).
The meat is 25% fat (negative frame).

The positively-framed quantity is consistently evaluated as more favourable than the complementary negative equivalent (Donovan & Jalleh, 1999; Kim et al., 2014; Levin & Gaeth, 1988; Seta et al., 2010). Past research suggests that changing the presentation format of the quantifier (e.g., from numerical to verbal) can affect the size of the framing effect (Welkenhuysen et al., 2001). For example:

The meat has a high [75]% of lean meat (positive frame).
The meat has a moderate [25]% of fat meat (negative frame).

The verbal quantifier format was found to magnify the framing effect (Welkenhuysen et al., 2001). However, this effect needs to be replicated. Further, two proposed theoretical accounts of attribute framing could explain why quantifier format should moderate the attribute framing effect, but the assumptions of these theories need to be empirically tested. In this paper, we build on past work by testing whether quantifier format (verbal or numerical) moderates the attribute framing effect, and investigating two possible explanations for such moderation.

7.2.1 Empirical tests of quantifier format as a moderator of attribute framing

Based on past empirical work, we could expect quantifier format to moderate the attribute framing effect. Although the effect is predominantly studied with complementary numbers and opposing attributes (e.g., 20% fat vs. 80% lean), some work shows that framing is also possible with verbal quantifiers (Reyna & Brainerd, 1991; Welkenhuysen et al., 2001). However, this work is limited. Reyna & Brainerd (1991)'s study focused on risky choice framing, which shows that people reverse their preferences between a sure and uncertain option when the frames are switched. This is a different type of framing effect from attribute framing, which compares two descriptions of the same item (Levin et al., 1998). To our knowledge, only one study has compared numerical and verbal quantifiers directly in the attribute framing effect (Welkenhuysen et al., 2001). In this study, participants were more likely to prefer a prenatal test when informed they had a moderate chance of having a baby with cystic fibrosis (negative frame) than a high chance of having a baby without (positive frame); however, they did not show this preference when given a 25% chance of cystic fibrosis vs. 75% chance of no cystic fibrosis (Welkenhuysen et al., 2001).

Three limitations of this research constrain the conclusion that format moderates the framing effect. First, this moderation effect was limited to one context and a set of materials. Hence, it has to be conceptually replicated in a different context and with a different set of materials. Second, individual variability in the interpretation of verbal quantifiers (e.g., Budescu et al., 2014) means that averagely-translated verbal quantifier frames may not be truly equivalent to an individual. For example, a moderate chance may not be perceived as the logical complementary equivalent to a high chance. Finally, variability in interpreting verbal quantifiers may be asymmetric over different quantifier magnitudes (Juanchich et al., 2019). Our own work in this domain has found that the meaning of high might vary between 30-80% across individuals, but the meaning of moderate might only vary between 20-50% (see Chapter 2). This could mean that the larger verbal framing effect could be a product of more difference in magnitude between the verbal frames than the numerical ones. A stringent test of the quantifier moderation effect on attribute framing would thus need to account for individual variability in translation, and also the magnitude of the quantifier across frames.

7.2.2 Theoretical rationale for quantifier format as a moderator of attribute framing

Although there is limited empirical evidence for quantifier format as a moderator of attribute framing, there is a theoretical basis to expect that changing the format would affect the size of the attribute framing effect. Two accounts of attribute framing suggest that verbal quantifiers should magnify the attribute framing effect compared to numerical quantifiers, albeit for different reasons.

Verbal quantifiers increase the affective response to a frame. The affective encoding account describes attribute framing as the product of an intuitive response primed by a frame's valence (Levin & Gaeth, 1988): a 75% lean meat is perceived as more positive than a 25% fat one because the positive affect associated with 'lean' primes people to judge the meat more favourably than the negative affect associated with 'fat'. Because verbal quantifiers are associated with more biases in judgement and decision-making than numerical ones, they are believed to encourage more intuitive processing than numerical ones (Windschitl & Wells, 1996). Hence, verbal quantifiers would be expected to magnify the affective response to the frame. This line of reasoning is supported by work that showed people used intuitive strategies that rely more on the valence of irrelevant contextual information for decisions with verbal vs. numerical quantifiers (see Chapter 4). Based on this evidence, we would expect people to rely more on the positive or negative affect of a verbal frame than a numerical one, and this should influence their judgement more, magnifying the framing effect.

Verbal quantifiers give more pragmatic signals about the speaker's recommendation. An alternative to the affective encoding account is the pragmatic inference account, which stresses that people infer implicit information from a speaker's choice of frame (Sher & McKenzie, 2008). One could reasonably infer that a speaker would choose to describe meat they recommended positively: as 'lean', rather than 'fat' (van Buiten & Keren, 2009). Verbal quantifiers are believed to carry more pragmatic meaning than numerical ones (Teigen & Brun, 2003). For example, people believed more that a speaker was being polite about an uncertain event when verbal probabilities were used compared to numerical ones (Sirota & Juanchich, 2012). Based on this reasoning, verbal quantifiers would allow people to make stronger inferences about the communicator's recommendations (Hilton, 2008; Keren, 2007) than numerical quantifiers.

7.2.3 Understanding the mechanisms involved in a moderated framing effect

While the affective encoding and pragmatic inference accounts predict the same result —a larger verbal than numerical attribute framing effect —they posit different mechanisms for attribute framing, and therefore different explanations for why the verbal format should magnify the effect. According to the affective encoding account, verbal formats magnify the framing effect by bolstering the affect encoded when processing the frame (Windschitl & Wells, 1996). According to the pragmatic inference account, verbal formats magnify the framing effect by bolstering the inferences drawn from the frame (Hilton, 2008; Keren, 2007). We sought to test these two explanations in parallel to assess how much affect or pragmatic inference contributes to the attribute framing effect, and why format moderates it.

It is important to note while the pragmatic inference account could exclude affect (one could, for instance, make an inference about a frame choice which may not involve affect, e.g., 'half-full' vs. 'half-empty'; Ingram et al., 2014), it does not necessarily rule out affect as a contributor to attribute framing. The account simply entails that the choice of valence offers a clue about the speaker's intentions, and it is this inferred intention, which may encompass some affect associated with the valence choice, that explains the framing effect. In contrast, affective encoding relies on automatic emotion priming (Levin, 1987)and should thus occur in conjunction with, or before a pragmatic inference, but not after (Gawronski & Ye, 2014). Following up on this, a key difference between the accounts lies in the way they conceptualise the framing effect: according to the pragmatic account, forming a favourable judgement of the 75% lean vs. 25% fat meat is rational as the recipient is using implied advice from the speaker to inform their judgement (Sher & McKenzie, 2008). Conversely, the affective encoding account entails that such discrepancy in one's judgement approximates to an irrational bias (Tversky & Kahneman, 1986).

7.2.4 The present work

Our research goals were to test whether verbal quantifiers would create a larger attribute framing effect than numerical quantifiers. This hypothesis was driven by previous empirical work, and reasoning from two empirical accounts: affective encoding and pragmatic inference. According to these accounts, either affect or inferences are involved in producing the attribute framing effect. The affective encoding account predicts that affect should mediate the framing effect, while the pragmatic inference account predicts that perceiving a speaker to recommend the item should mediate the framing effect. Therefore, we derived two novel hypotheses: based on the affective encoding account, format would moderate a framing effect mediated by affect, but based on the pragmatic inference account, format would moderate a framing effect mediated by inferred recommendations. These hypotheses are illustrated in Figure 7.1.

We report the results of four studies that systematically tested our hypotheses. In Experiments 1 and 2, we tested whether format moderated the framing effect across three quantifiers of different magnitudes (low, medium, and high), and if this moderation occurred via the affect associated with the valence of the frame. We used numerical quantities known to be average translations of the verbal quantities (Experiment 1) and numerical quantities produced by participants themselves as translations of the verbal quantities. (Experiment 2). In Experiment 3, we tested the role of inferred recommendations from a speaker in accounting for the moderated framing effect. Experiment 4 compared the two explanations in a combined model and tested for affect and pragmatic inferences

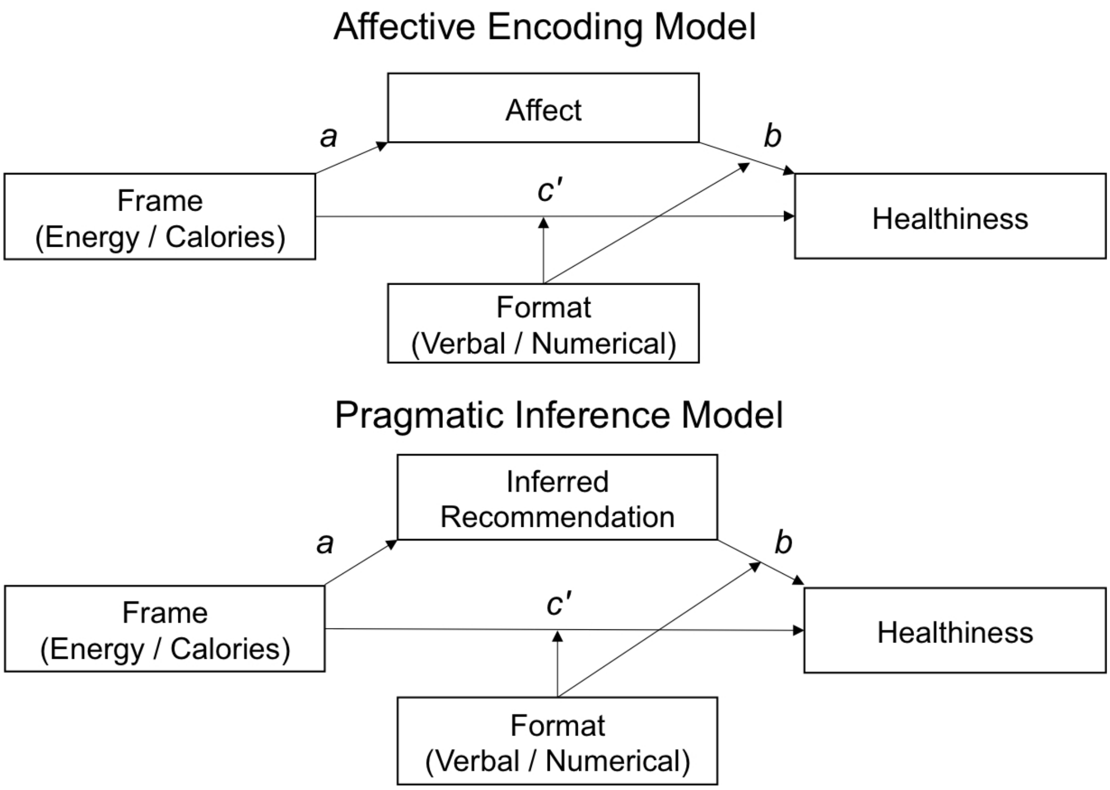

Figure 7.1. Moderated mediation models tested to investigate predictions of the affective encoding account (Experiments 1 & 2) and the pragmatic inference accounts (Experiment 3). Both accounts were tested in parallel in Experiment 4 (see Figure 7.4).

as parallel mediators.

7.2.5 Open science statement

In line with recent scientific guidelines, the methods and analyses for all four experiments were pre-registered, and can be found on the Open Science Framework along with the materials and data.

7.3 Experiment 1

7.3.1 Method

Participants. One hundred and ninety participants (62% female; 71% White; ages 18-79, M = 32.51, SD = 14.14) were sourced from a university lab database. The sample size was determined *a priori* based on a stopping rule that specified a specific data for data collection to cease. No analyses were performed prior to the completion of data collection. On average, participants' BMI was in a slightly overweight category (M = 25.92, SD = 10.13) and positive attitudes towards healthy eating (Steptoe et al., 1995; M = 5.14, SD = 1.02, on a 7-point scale). Sixty-seven percent used nutrition labels frequently.

Design. Participants completed the experiment on Qualtrics at the end of an unrelated 20-minute medical survey. Participants were randomly assigned across four conditions in a 2 (format: verbal or numerical; between-subjects) × 2 (frame: positive or negative; between-subjects) × 3 (quantity magnitude: low/20%, medium/40%, or high/70%; within-subjects) mixed design.

Materials and procedure. To ensure the equivalence of both frames, we used a novel framing design that kept the quantity stable across the attribute framing conditions. This allowed us to manipulate both attribute and quantity orthogonally and disentangle their respective roles. We used the domain of food energy, where in the context of food, people view 'energy' positively, but the interchangeable term 'calories' negatively (Watson et al., 2013). This is akin to framing effects based on the goal framing approach (Epley et al., 2006; Gamliel, 2013; Hardisty et al., 2010).

In our framing scenario, participants read the following vignette:

A food product is labelled with the following information:
'Provides [quantity %] of your daily [attribute].'

The attribute was either energy or calories. The quantity was given either in verbal or numerical format. Participants read the same vignette with three quantities: low (20%), medium (40%), or high (70%), presented in random order. Participants rated the healthiness of the food on a 7-point Likert scale (1: *very unhealthy*, 7: *very healthy*).

As a measure of affective associations with the attribute, participants next rated the terms 'energy' and 'calories' individually on a 9-point semantic differential scale with four sets of bipolar adjectives (e.g., bad –good; MacGregor et al., 2000). Participants also judged their affective associations with eight filler nutrients, such as ‘protein’ or ‘sugar’. Scale reliability was good, Cronbach’s α= .95 (energy) and .94 (calories). Affect was calculated as the mean for the four adjectives on the scale, with higher scores indicating a more positive affective association.

Finally, participants reported their attitudes towards healthy eating (Steptoe et al., 1995), how frequently they used nutrition labels, and socio-demographic information, including weight and height.

7.3.2 Results

Analysis strategy. The model tested for Experiment 1 is illustrated in the top panel of Figure 1. We tested this moderated mediation model on each quantity (small, medium, and large), using the PROCESS macro in SPSS (Model 15, using bias-corrected bootstrap confidence intervals with 5,000 samples; Hayes, 2013). This is a conditional process model that assesses direct and indirect effects, along with mediation and moderation. According to the affective encoding hypothesis, we expected an effect of frame on healthiness to be mediated through affect (indirect effect), with format moderating the path between affect and healthiness.

Framing effect on healthiness judgements. The pattern of healthiness judgements for small and large quantities is shown in the top panel of Figure 7.2 (medium quantities had an almost identical, but not significant, pattern to large quantities; we provide the graph for medium quantities in Appendix F). Overall, participants exhibited a framing effect, albeit not always in the predicted direction. For small quantities, they judged small (low/20%) calories as healthier than the equivalent energy, b = -0.83, p = .001. For the other quantities, they judged medium (40%) and large (high/70%) energy as healthier than the equivalent calories, but this was only significant for the large quantities, b = 0.44, p = .088 (medium), b = 1.48, p < .001 (large).

Testing for format moderation of the framing effect. The difference in magnitude of the framing effect between formats was in the expected direction (greater for verbal than numerical quantifiers), but the interaction effects that would quantify the moderation were not significant, b = 0.64, p = .083

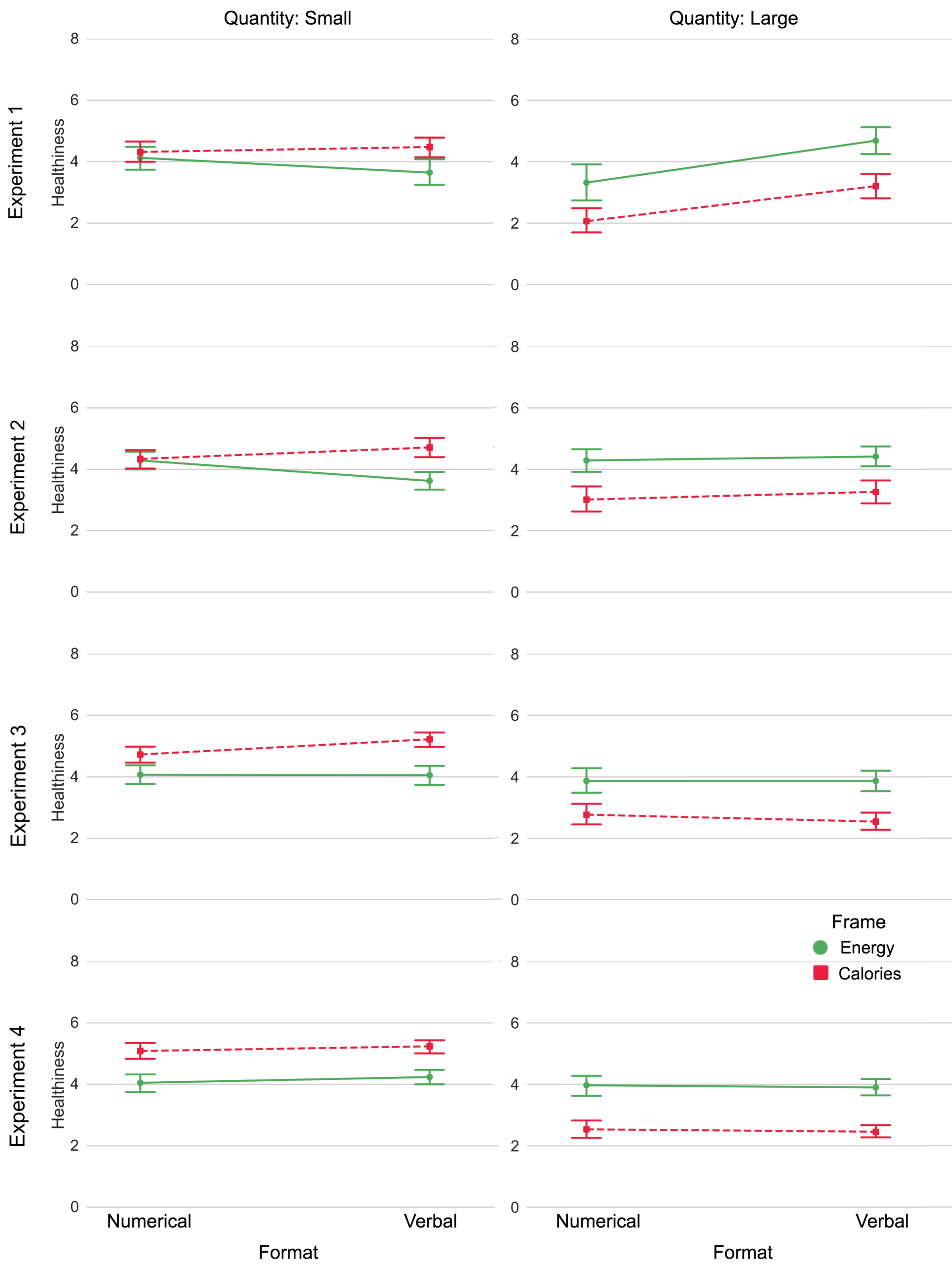

Figure 7.2. Effects of frame, format, and quantity (small and large) on healthiness judgements across Experiments 1-4, with error bars reflecting 95% confidence intervals. The difference between the green solid line and the red dotted line shows the magnitude of the framing effect. We expected the difference to be larger for verbal than numerical quantifiers.

Note. Medium quantities are omitted from the figure as they were only tested in Experiment 1. The corresponding graph for medium quantities in Experiment 1 is provided in Appendix F.

(small), b = 0.12, p = .750 (medium), b = -0.22, p = .642 (large).

Does affect explain the framing effect? Although we did not find significant evidence of the prediction moderation by format, we still tested as planned whether affect predicted variations in the magnitude of the framing effect. We first checked whether affect was independently correlated with frames and healthiness judgement. As illustrated in the top left panel of Figure 7.3, participants associated more positive affect with energy than calories (on the 9-point semantic differential scale: (M_{energy} = 6.00, SD = 1.04; M_{calories} = 4.62, SD = 1.51), b = 1.38, $p < .001$. Affect predicted healthiness for both the small and large quantities, however this effect was in opposite directions (negative for small, positive for large), b = -0.25, p = .002 (small), b = 0.31, p = .001 (large). Affect did not predict healthiness for medium (40%) quantities, b = 0.11, p = .135. Given that affect only predicted healthiness judgement for low and high quantities, we only report below the mediation effects for these models. A full report of the model coefficients for all quantities is in Appendix F (Table F.1).

Mediation by affect. The conditional indirect effect of frame on healthiness assessed the extent to which affect mediated the framing effect on healthiness judgements for each format. For small quantities, affect significantly mediated the frame-healthiness relationship for both verbal and numerical quantities, b = -0.35, 95% CI [-0.71, -0.01] (verbal); b = -0.33, 95% CI [-0.64, -0.07] (numerical). For large quantities, affect significantly mediated the frame-healthiness relationship for verbal, but not numerical, quantities, b = 0.60, 95% CI [0.26, 1.06] (verbal); b = 0.26, 95% CI [-0.07, 0.60] (numerical).

Framing effect after accounting for affect as a mediator. The conditional direct effect of frame on healthiness measured how the framing manipulation predicted healthiness judgement independently of the mediator, affect. For small quantities, we found a trend of larger direct effect with the verbal than numerical quantifiers, but this interaction did not reach significance, b = 0.76, p = .076. For large quantities, frame had a significant direct effect on healthiness, which did not vary according to format, b = 0.06, p = .898.

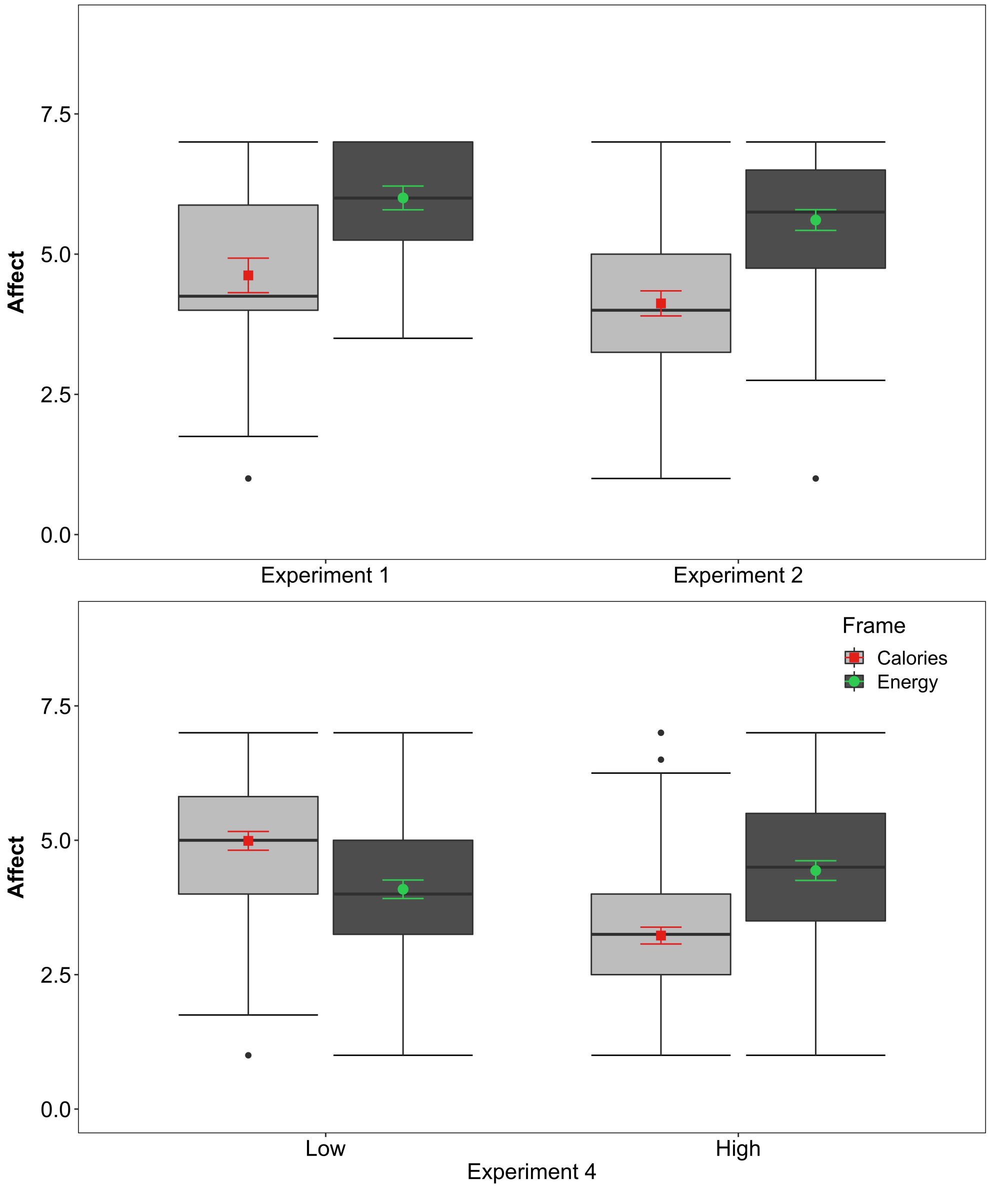

Figure 7.3. Affect rating for energy and calorie frames in Experiments 1, 2, and 4 on a 9-point semantic differential scale. The attribute frames were rated in isolation in Experiments 1 and 2, and together with the quantifier 'low' or 'high' in Experiment 4.

Note. The point plots (red and green) give the means and 95% confidence intervals of participants' responses, and the box-and-whisker plots show the overall distributions of the responses. All differences in affect ratings between frames were significant.

7.3.3 Discussion

Experiment 1 showed that format and affect had complex roles in the framing effect. In contrast to our expectations, format did not moderate the framing effect. However, we also found evidence that format had an effect on how affect mediated the framing effect, contingent on which quantity participants judged: the mediation via affect was significant for the large verbal but not large numerical quantities. Further, our results showed an interesting reversal of judgements between small and large quantities. The healthiness ratings for energy quantities (positive frame) were similar, but small quantities of calories (the negative frame) were judged as healthier than energy whereas large quantities of calories were judged as less healthy than energy. This result brings a new perspective to attribute framing, because unlike in traditional attribute framing studies, which compare a small negative attribute with a large positive one (or vice versa), the magnitude of the quantity was independent from the frame.

However, before drawing a conclusion from Experiment 1 about the roles of format and affect, we needed to consider a constraint on interpretation of the observed differences in affect and healthiness judgements. Although we selected verbal and numerical percentages for energy that are on average psychologically equivalent (see Chapter 2), individual participants may not have perceived them as such (Berry et al., 2002; Budescu et al., 2014). Participants could, for example, believe that 'low calories' meant less than 20%, which would make the frames inequivalent between formats. We thus sought to replicate our results while controlling fro this individual variability.

7.4 Experiment 2

The goal of Experiment 2 was to test the role of format and affect as moderator and mediator of the attribute framing effect in an improved design that accounted for the variability in interpretations across verbal and numerical quantifiers. To do so, we had participants first translate the verbal percentages of energy or calories into a numerical percentage, and their translations were later used to solicit their healthiness judgements. Experiment 2 additionally investigated whether participants' healthiness judgements would correspond with their willingness to pay for the food.

7.4.1 Method

Participants. Participants were sourced from a survey panel (*N* = 194; rewarded with online vouchers) and undergraduate participant pool (N = 141; rewarded with course credit)[1]. The experiment was powered to detect a frame × format interaction based on the effect size observed in Experiment 1 (f = .10, α= .05, 1-β= .80). Participants were 59% female, 80% White, with ages ranging from 18-76 (*M* = 37.76, *SD* = 17.30). Mean BMI was healthy (*M* = 24.99, *SD* = 5.78). Participants' attitudes towards healthy eating were on average slightly positive (*M* = 4.89, *SD* = 0.98) and 53% used nutrition labels frequently.

Design. The design was the same as Experiment 1 (frame and format manipulated between-subjects; quantity magnitude manipulated within-subjects), quantity manipulated within-subjects), except that we used only two quantities, a small (low) and large (high) one.

Materials and procedure. Participants first provided numerical estimates for a low % and high % of either calories or energy[2]. These translations were used as the within-subject quantity manipulation in the numerical condition. As part of the translation task, participants also provided four filler numerical-to-verbal translations of other food quantities (e.g., 25% fat meat).

To distract participants from focusing on the translations they had provided, participants next completed a filler task similar to one used in Teigen et al. (2014). After this, participants completed the healthiness judgement task from Experiment 1 for the small and large quantities (verbal or numerical). Participants then indicated how much they would be willing to pay for a cereal bar with

[1]A multivariate ANOVA found no evidence that the two samples differed significantly in their response to the manipulations. All subsequent analyses were therefore performed on the combined sample.

[2]The average numerical translations were similar to the ones in Experiment 1 (low: 22.65%, high: 66.98%).

that energy (calorie) description (in pounds sterling). The judgement task also included four filler items corresponding to the fillers from the earlier translation task.

Finally, participants completed the semantic differential scale (MacGregor et al., 2000) for 'energy' or 'calories' and the same socio-demographic measures as in Experiment 1.

7.4.2 Results

Framing effect moderated by quantifier format. The pattern of results was similar to Experiment 1 (see second panel of Figure 7.2). The frame had an effect on perceived healthiness: participants judged small quantities of calories healthier than small quantities of energy, but large quantities of energy healthier than large quantities of calories, b = -1.09, p < .001 (small), b = 1.15, p < .001 (large). The framing effect for willingness-to-pay was not significant, b = -0.04, p = .843 (small); b = 0.27, p = .134 (large).

Testing for format moderation of the framing effect. Format moderated the attribute framing effect for small, but not large, quantities. Small verbal quantifiers produced a magnified framing effect compared to small numerical quantifiers, b = 1.04, p = .001. There was no significant interaction for large quantities, b = 0.13, p = .743. Given that framing did not affect willingness-to-pay, we do not follow up on the moderated mediation analysis for this variable here (but results can be found in Appendix F).

Does affect explain the framing effect? We tested the moderated mediation of the frame-healthiness relationship using the same conditional process model as used in Experiment 1 (see top panel of Figure 7.1; beta coefficients for all analyses reported in the Appendix F, Table F.2). As illustrated in the top right panel of Figure 7.3, energy evoked significantly more positive affect than calories (M_{energy} = 5.61, SD = 1.22; $M_{calories}$ = 4.12, SD = 1.46), b = -1.49, p < .001. In line with the affective encoding account, positive affect was significantly positively associated with both healthiness and willingness-to-pay for large quantities, b = 0.39, p < .001; b = 0.17, p = .001. However, affect did not directly lead to higher healthiness judgements or willingness-to-pay for small quantities, b = -0.09, p =

.235 (healthiness); b = -0.05, p = .325 (willingness-to-pay).

Mediation by affect. The conditional indirect effect of frame on healthiness showed the same pattern for both verbal and numerical quantifiers. For large quantities, we found a significant mediation by affect for both formats, b = 0.56, 95% CI [0.28, 0.89] (verbal); b = 0.60, 95% CI [0.28, 0.97] (numerical). However, the mediation was not significant for small quantities, b = -0.18, 95% CI [-0.47, 0.07] (verbal); b = -0.08, 95% CI [-0.42, 0.24] (numerical). The test for moderation of the indirect effect was not significant (index of moderated mediation$_{\text{small}}$ = .10, 95% CI [-.31, .54], index of moderated mediation$_{\text{large}}$ = .03, 95% CI [-.36, .46].

The analysis on willingness-to-pay showed only a significant mediation for large quantities, b = 0.23, 95% CI [0.06, 0.45] (verbal); b = 0.27, 95% CI [0.06, 0.52] (numerical); all other CIs straddling 0.

Framing effect after accounting for affect as a mediator. When we controlled for the role of affect, the frame still had an effect on judgements. We found a significant direct framing effect for both quantities, b = -0.43, p = .023 (small), b = 0.63, p = .006 (large). The direct effect (after accounting for affect) was significant only for verbal, but not numerical, quantifiers, b_{small} = -0.91, p < .001 (verbal), b_{small} = 0.03, p = .903 (numerical), b_{large} = 0.61, p = .038 (verbal), b_{large} = 0.66, p = .060 (numerical). It should be noted, however, that the interaction term in the direct effect (controlling for affect) was significant only for small and not large quantities, b_{small} = 0.95, p = .013, b_{large} = 0.05, p = .906.

7.4.3 Discussion

In Experiment 2, we aimed to re-test the hypotheses of Experiment 1 with an improved design that better controlled for inter-individual variability in the interpretation of verbal quantifiers. We used numerical quantifiers tailored to individual participants' interpretation of the verbal quantifiers. When we ensured equivalence between the verbal and numerical quantifier phrases, we observed a framing effect on healthiness judgements that was larger for small verbal than small numerical quantifiers, and reversed in direction between large and small quantities. However, the framing effect did not extend consistently to willingness-to-pay, perhaps because consumers do not necessarily prefer buying healthier products (Raghunathan et al., 2006).

The results of the mediation analysis indicated that people judged high energy as healthier than high calories because of the greater affect associated with energy than calories, in line with the predictions of the affective encoding account. However, affect did not mediate the framing effect for small quantities, although this effect was moderated by format. There are two possible explanations for this. First, a different mechanism might account for attribute framing, especially with small quantities. For example, the hypothesis presented by the pragmatic account of framing effects posits that people factor inferences they draw from a given frame into their judgements (Sher & McKenzie, 2008), and verbal quantifiers, with stronger pragmatic signals (Teigen & Brun, 2000), should increase these inferences. We tested this hypothesis in Experiment 3. Second, participants might have derived the affect not only from the frame, but from the combination of quantifier and frame (which was reversed in the small quantities). We followed up on this possibility in Experiment 4.

7.5 Experiment 3

The goal of Experiment 3 was to test whether quantifier format moderates the attribute framing effect through the mechanism of stronger pragmatic inferences, as posited by the pragmatic inference account. As speakers tend to use positive frames to recommend options (and vice versa; van Buiten & Keren, 2009), and listeners may infer speaker intentions upon receiving the communication (Keren, 2007), we expected participants to infer the speaker's recommendation based on the frame. On seeing a food with a large portion of energy, we expected participants would infer it to be more recommended than a food with a large portion of calories (vice versa for small portions). Since verbal quantifiers allow people to infer more pragmatic information than numerical ones (e.g., whether a speaker is being polite; Sirota & Juanchich, 2012), we hypothesised that participants would infer stronger recommendations coming from verbal quantifiers than numerical ones.

7.5.1 Method

Participants. We sourced participants through a university sample (N = 71; participation voluntary) and a survey panel (N = 223; paid £0.70)[3]. Participants were 69% female, 88% White, with ages ranging from 18-71 (M = 37.38, SD = 13.24). Participants' BMI was on average in the overweight category (M = 27.03, SD = 7.18). Their attitudes towards healthy eating were slightly positive (Steptoe et al., 1995; M = 4.94, SD = 0.95). Seventy-one percent used nutrition labels frequently.

Design. The design was identical to Experiment 2 (2: frame × 2: format × 2: quantity magnitude, mixed design, quantity magnitude within-subjects).

Materials and procedure. Participants provided numerical translations of verbal quantities and completed a filler task before completing the healthiness judgement task. After this, they rated how much they agreed with three statements about what the speaker believed the information recipient should do about the food product: that they should buy the food, that the food was healthy, and that the product was good. Participants provided this rating on a 7-point Likert scale (1: *strongly disagree*, 7: *strongly agree*). Scale reliability was good when applied to both small and large quantities, Cronbach's α= .89 (small) and .86 (large). We computed an inferred recommendation rating from the average of the three scale measures. Finally, we collected eating attitude and socio-demographic measures as in the previous experiments.

7.5.2 Results

Analysis strategy. We tested the moderated mediation model on each quantity (small and large) using Model 8 in the SPSS PROCESS macro (see bottom panel of Figure 7.1; Hayes, 2013). As we expected that format would moderate the effect of frame on both healthiness and strength of inferred recommendations, we used this model to include a moderated pathway between frame

[3] A multivariate ANOVA found no evidence that the two samples differed significantly in their response to the manipulations. All subsequent analyses were performed on the combined sample.

and recommendations. Full beta coefficients for the analyses are reported in Appendix F (Table F.3).

Framing effect on healthiness judgements and inferred recommendations. We again produced a framing effect, with participants finding small quantities of calories healthier than energy, but large quantities of energy healthier than calories (see third panel of Figure 7.2), b = -1.17, p < .001 (small), b = 1.33, p < .001 (large). Although the framing effect was slightly greater for the verbal than numerical quantifiers, this moderation was not significant, b = 0.51, p = .078 (small), b = -0.23, p = .520 (large).

Testing for format moderation of the framing effect. Contrary to expectations, we did not find the predicted moderating effect of format on healthiness judgements or inferred recommendations. As illustrated in Figure 7.2, the framing effect on healthiness judgements was slightly greater for verbal than numerical quantifiers, but the interaction was not significant, b = 0.51, p = .078 (small), b = -0.23, p = .520 (large). The interaction was also not significant for inferred recommendations, b = 0.48, p = .069 (small), b = -0.42, p = .135 (large).

Do inferred recommendations explain the framing effect? We tested using the planned analyses whether inferred recommendations mediated the effect of frame on healthiness judgements.

Mediation by inferred recommendations. Inferred recommendations significantly mediated the frame-healthiness relationship. This was the case for both quantities and both formats, b_{verbal} = -0.67, 95% CI [-.95, -.43]; $b_{\text{numerical}}$ = -0.36, 95% CI [-.62, -.13] (small), b_{verbal} = 0.82, 95% CI [.51, 1.14]; $b_{\text{numerical}}$ = 0.50, 95% CI [.19, .84] (large). We found no evidence of a moderated mediation (index of moderated mediation$_{\text{small}}$ = .31, 95% CI [-.02, .65], index of moderated mediation$_{\text{large}}$ = -.32, 95% CI [-.76, .07].

Framing effect after accounting for inferred recommendations as a mediator. After controlling for inferred recommendations, we still found a significant direct framing effect, b = -0.40, p = .002 (small), b = 0.56, p < .001 (large). This was not significantly moderated by format, b = 0.20, p = .394 (small), b = 0.09, p = .750 (large).

7.5.3 Discussion

Contrary to our hypothesis, verbal quantifiers did not significantly magnify the framing effect in Experiment 3. However, there was a significant mediation that suggested participants judged a product to be healthier because they believed it to be more recommended. This could indicate that while pragmatic inference plays a role in framed judgements, the pragmatic advantage of verbal over numerical quantifiers may not be as great as previously suggested (Teigen & Brun, 1995), especially when interpretational variability is taken into account.

Nonetheless, through Experiments 1 to 3, we observed independent empirical support for both affective and pragmatic accounts as explanations for the framing effect: why the same quantity of energy and calories are judged so differently in healthiness. In Experiment 4, we sought to provide a final high-powered test of the moderated framing effect, and integrate both affective encoding and pragmatic inference accounts in a combined model and compare their relative contributions to the attribute framing effect.

7.6 Experiment 4

Experiment 4 had three goals. First, we aimed to test the moderated framing effect: whether verbal quantifiers would magnify the effect. Second, we tested both affect and inferred recommendations in parallel to compare their relative contributions as mediators of the attribute framing effect (with and without format moderation). Third, we followed up for the possibility from Experiment 2 that small quantities reversed the affect of the frame. We controlled for this by measuring the affect associated with the entire quantified phrase (e.g., 'low calories') rather than just the independent attribute (i.e., 'calories').

7.6.1 Method

Participants. The experiment was powered to detect the smallest mediation effect size obtained in the previous experiments, which was for affect in Experiment 2 (α path = .49, β path = .16; Fritz & MacKinnon, 2007). Four hundred and eight participants (68% female, 28% male, 4% undisclosed; 86% White, age range 18-85, M = 35.54, SD = 12.12) were offered £0.70 to complete the experiment on Prolific Academic. Participants' average BMI was in the overweight category (M = 26.94, SD = 7.73). Their attitudes towards healthy eating were slightly positive (Steptoe et al., 1995; M = 4.91, SD = 0.99). Seventy-three percent used nutrition labels frequently.

Design. The design was the same as Experiments 2 and 3: frame (energy or calories) and format (verbal or numerical) manipulated between-subjects, quantity magnitude (small and large) within-subjects.

Materials and procedure. Participants provided numerical translations of verbal quantifiers, followed by a filler task. Participants then completed in random order: the healthiness judgement task (same as Experiments 1-3), the inferred recommendation scale (same as Experiment 3; Cronbach's α_{small} = .92, α_{large} = .91), and the semantic differential scale for energy or calories (MacGregor et al., 2000; Cronbach's α_{small} = .90, α_{large} = .87). Participants completed each of these three tasks twice in randomised order, once for the small and once for the large quantifier. Finally, eating attitudes and socio-demographic measures were collected.

7.6.2 Results

Analysis strategy. To assess the moderated framing effect, we first tested a moderated mediation model with parallel mediators at each quantity level (PROCESS Model 8; Hayes, 2013; see top panel of Figure 7.4). Given the possibility that format would not significantly moderate the effect (based on Experiment 3), we also planned a parallel mediation model without moderators (PROCESS Model 4; Hayes, 2013); see bottom panel of Figure 3). We opted for parallel mediators because the theoretical accounts behind the explanatory mechanisms we tested do not specify an interactive role between mediators.

Framing effect on healthiness judgements, affect, and inferred recommendation. As shown in the bottom panel of Figure 7.2, the framing effect reversed depending on the quantity of calories: small quantities of calories were more healthy than the same amount of energy, but large quantities of calories were more healthy than the same amount of energy, b = -1.00, $p < .001$ (small),

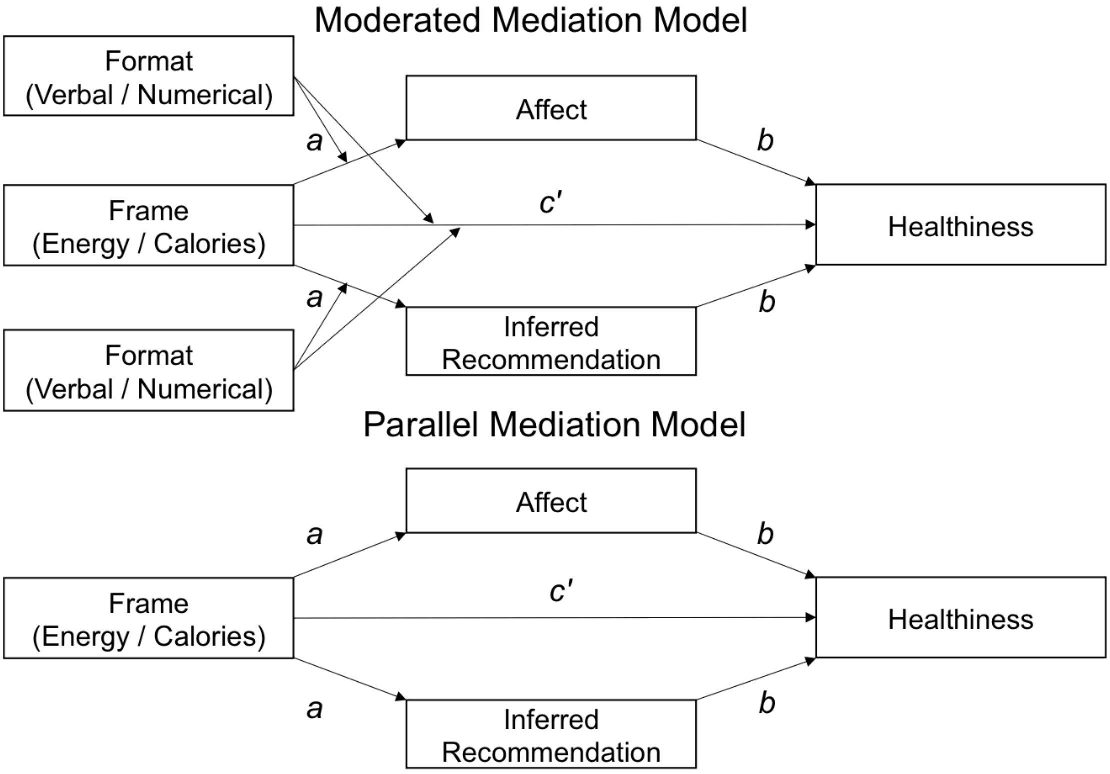

Figure 7.4. Mediation models tested in Experiment 4, comparing two mediators working in parallel (with and without moderators). The moderated mediation did not find a significant moderation by format.

b = 1.44, p < .001 (large). This pattern was consistent for the affect associated with the frame, as illustrated in the bottom panel of Figure 7.3 —while affect was similar for different quantities of energy, it was higher for low calories than high calories, b = -0.90, p < .001 (small), b = 1.21, p < .001 (large). It was also consistent for inferred recommendations, b = -0.94, p < .001 (small), b = 1.15, p < .001 (large).

Testing for format moderation of the framing effect. Format did not moderate the framing effect on healthiness, b = -0.04, p = .886 (small), b = -0.01, p = .978 (large). There was also no significant frame format interaction for affect or inferred recommendations, $b_{\text{small}} = 0.07$, p = .773, $b_{\text{large}} = 0.25$, p = .324 (affect), $b_{\text{small}} = 0.26$, p = .267, $b_{\text{large}} = 0.10$, p = .702 (inferred recommendations).

Do affect and inferred recommendations explain the framing effect? Both affect and inferred recommendation mediated the frame-healthiness relationship, but format did not moderate either the direct or indirect effects (all

CIs for the interaction straddling 0; full beta coefficients reported in Appendix F, Table F.4). We therefore performed the planned parallel mediation analysis with affect and inferred recommendation without moderators. Table 7.1 reports the beta coefficients for the separate direct and indirect effects, showing a partial mediation for both mediators. In addition, the parallel mediation analysis also tested for the difference in the two indirect effects (mediation by affect vs. mediation by inferred recommendations). This comparative test of the indirect effects for the two mediators was not significant at small quantities, indicating that both mediators were working in parallel, $b = .12$, 95% CI [-.10, .35]. However, there was a significantly larger indirect effect (stronger mediation) for affect than inferred recommendations for large quantities, $b = -.40$, 95% CI [-.68, -.11]. This indicated that the affect associated with large quantities better explained why people found energy healthier than calories, as opposed to whether they inferred the speaker to recommend the large quantities in the food.

7.6.3 Comparison of format moderation across experiments

Experiment 4 tested whether format (verbal vs. numerical) moderated the framing effect. Contrasting Experiment 2, but in line with Experiments 1 and 3, we did not find that format significantly moderated the effect of framing for healthiness judgement, affect, or inferred recommendations. Given the mixed results on the effect of quantifier format, we tested for the robustness of the moderation by meta-analysing the framing effect across experiments as moderated by format and quantity magnitude. A meta-analysis allowed us to estimate the moderating effects of format and quantity magnitude with greater precision and statistical power (Cumming, 2013). We used the R package 'metafor' (Viechtbauer, 2010) to meta-analyse our four experiments using random effect models (a restricted maximum likelihood estimator).

The effect of format was significantly different between quantity magnitudes (large and small), as evidenced by a significant interaction between the frame and quantity magnitude moderators in the meta-analysis (see Figure 7.5), $b = -0.63$, $p = .005$, 95% CI [-1.07, -0.19]. We thus followed up with analyses for the small and large quantities respectively. The moderating effect of format was

Table 7.1. Direct and indirect effects of frame on affect, inferred recommendation, and healthiness judgement in Experiment 4 (PROCESS Model 4).

Factors	***b***	**SE_b**	***t***	***p***
Effect of frame on affect (a path)				
Small quantities (R_2 = .11, p < .001)	-0.90	.12	-7.23	< .001
Large quantities (R_2 = .20, p < .001)	1.21	.12	9.87	< .001
Effect of frame on recommendation (a path)				
Small quantities (R_2 = .14, p < .001)	-0.94	.12	-8.07	< .001
Large quantities (R_2 = .17, p < .001)	1.15	.13	9.12	< .001
Effects of frame, affect, and recommendation on healthiness				
For small quantities (R_2 = .53, p < .001):				
Frame (c' path)	-0.29	.10	-2.78	.006
Affect (b path)	0.47	.06	8.03	< .001
Recommendation (b path)	0.33	.06	5.69	< .001
For large quantities (R_2 = .53, p < .001):				
Frame (c' path)	0.28	.11	2.46	.014
Affect (b path)	0.64	.07	9.70	< .001
Recommendation (b path)	0.33	.58	5.71	< .001

Note. The bottom panel of Figure 7.4 illustrates the corresponding pathways for each of the effects in the parallel mediation model. The a path is the direct effect of frame on the respective mediators (1: affect, 2: inferred recommendation); the b path is the direct effect of each of the mediators on healthiness; the c' path is the direct effect of frame on healthiness after accounting for the mediated pathways.

significant for small quantities, with verbal quantifiers magnifying the framing effect across studies, b = 0.43, p = .023, 95% CI [0.06, 0.79]. However, format did not moderate the effect significantly for large quantities, b = -0.21, p = .098, 95% CI [-0.45, 0.04].

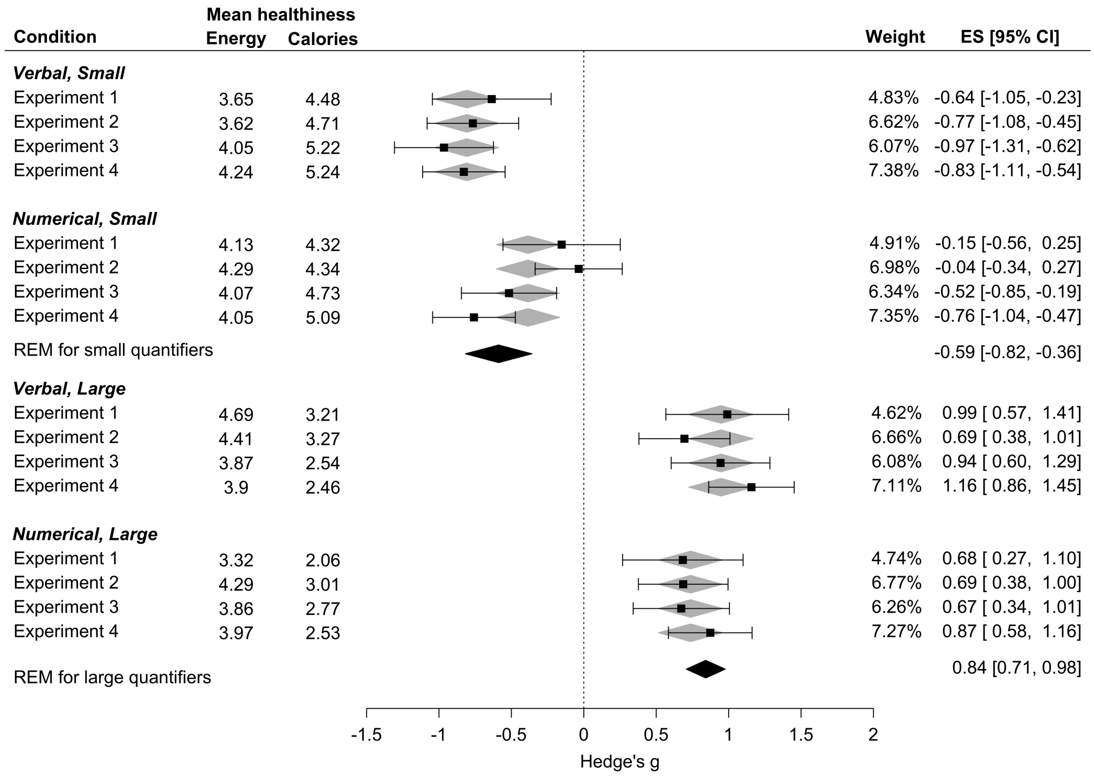

Figure 7.5. Forest plot showing the framing effect size (energy –calories) across four experiments for the verbal and numerical small and large quantities.

Note. Grey diamonds show the overall effect within the format subgroup for each quantity magnitude. Black diamonds show the random effect model for small and large quantifiers respectively.

7.7 General Discussion

Over four pre-registered experiments, we investigated whether quantifier format (verbal vs. numerical) moderates the effect of attribute framing across quantity magnitudes. We also tested whether verbal formats would magnify the effect by increasing the intuitive response to the frame's affect (based on the affective encoding account; Levin & Gaeth, 1988), or by strengthening the inferences made about a speaker's recommendation for the framed item (based on the pragmatic inference account; Sher & McKenzie, 2008). Three main findings summarise our empirical results. First, we produced a robust attribute framing effect by varying only the attribute valence ('energy' vs. 'calories': relative to energy, the negative framing of calories resulted in food being judged significantly healthier or less healthy). However, we could only find evidence that the framing effect size was increased with verbal formats for small quantities when we analysed results across all four experiments. Second, both affect and inferred recommendations mediated the framing effect, highlighting the importance of combining theoretical explanations such as the affective encoding account and pragmatic inference account to explain psychological effects. Third, the direction of the framing effect was dependent on the magnitude of the quantifier and not simply the attribute, which suggests that the quantifier's magnitude modified the attribute's valence. This raises methodological considerations about how to control for this factor in future research.

7.7.1 Verbal quantifiers did not consistently magnify the framing effect

We expected to replicate previous work in which verbal framing effects were larger than numerical ones (Welkenhuysen et al., 2001), but this hypothesis was not convincingly supported. Our meta-analysis of the effect of format on the framing effect size found that across the four experiments, verbal quantifiers only magnified the framing effect compared to numerical quantifiers for small, but not large quantities. Two possibilities could account for these different findings. First, where previous research used verbal and numerical quantifiers that were translated as equivalent on average (Welkenhuysen et al., 2001), we used an individual translation paradigm that allowed us to control for variations in interpretation of verbal quantifiers. This ensured that observed differences between verbal and numerical quantifiers were not simply due to people interpreting the verbal quantifiers as having a different magnitude from the numerical ones. Second, our framing paradigm varied only the attribute between positive and negative frames, as opposed to traditional framing, where quantities are modified with the attribute to create complementary frames (e.g., 25% fat vs. 75% lean; Levin & Gaeth, 1988). In a previous study, we found that framing effect sizes can vary depending on the combination of quantities used (e.g., 5% fat or 95% lean produced a smaller framing effect than 25% fat or 75% lean), and this was associated with whether the verbal and numerical quantifier directed a reader's focus to the attribute (e.g., 5% fat prompted readers to think of lean meat, but 25% fat prompted readers to think of fat meat; see Chapter 6). Past findings that verbal quantifiers magnified the framing effect compared to numerical ones may not have accounted for the greater asymmetry between verbal and numerical complementary frames. When varying only the attribute while keeping the quantifier constant across frames, we showed that format only moderated small quantities.

7.7.2 Combining explanations for attribute framing

We used mediation analyses to test two explanations for the framing effect and variations in its size. These analyses found that people's affective associations and inferences about the speaker's recommendations both explained the effect of frame (i.e., whether the food was described in terms of energy or calories) on healthiness judgements. These findings fit two accounts of the framing effect that present different perspectives of the effect. The affective encoding account classifies the framing effect as an irrational bias (Tversky & Kahneman, 1986), primed by the encoding of positive or negative affect associated with the frame (Levin & Gaeth, 1988). In contrast, the pragmatic inference account views the framing effect as a product of rational reasoning (Sher & McKenzie, 2008), because people infer that the speaker chose a positive or negative frame to convey their preferences and recommendations about the target (Hilton, 2008; van Buiten & Keren, 2009). By testing affect and inferred recommendations as parallel mediators, we were able to compare the two mechanisms. Extending previous research, we found empirical evidence to support two mediation pathways. Compared to negative frames, positive frames were both associated with more positive affect and greater inferred recommendations from the speaker, and both mediators could account in tandem for the framing effect. As such, we suggest that explanations of the framing effect need to integrate both the affective encoding and pragmatic accounts.

One important limitation of mediation analysis is that it relies on correlations between the dependent variables (affect, inference, and healthiness), and thus cannot conclude that affect and inference have a causal effect on perceived healthiness. However, it is unlikely that the healthiness judgement preceded affect, because affect is typically automatic and instinctive (Murphy & Zajonc, 1993). In the case of pragmatic inferences, it is more debatable whether the inferred recommendation led people to believe the food was healthy or vice versa, especially as it is not certain whether pragmatic inferences are automatic or effortful (De Neys & Schaeken, 2007; Zhao et al., 2015). One way to disentangle this in future work may be to manipulate the ability to infer recommendations by presenting participants with the speaker's actual views along with the frame, and testing whether this changes inferences and subsequent healthiness judgements.

Another possibility to consider is how affect might be related to pragmatic inferences: the strength of the affect associated with the quantified frame could be a signal for how much to rely on pragmatic inferences to reach a judgement. We determined that both affect and inferred recommendations are drivers in the framing effect, however our selected mediation model, driven by our objective of comparing the accounts, did not account for a relationship between the mediators. Nonetheless, our results support an integrated and updated explanation for the framing effect: affective encoding and pragmatic inferences both contribute to people's processing of the combination of quantifier and attribute in a frame. Future work is still needed to adapt and integrate the two accounts into a combined theoretical model, and formally test the relationship between affect and inferred recommendations.

7.7.3 Framing effects match the valence of the quantified phrase, not just the attribute

While we observed a robust framing effect between energy and calorie frames across all quantities, the direction of participants' judgement differences reversed depending on whether the attribute had a large or small quantity. In contrast to previous attribute framing research, in which different quantities are attached to each frame (e.g., 75% lean vs. 25% fat; Levin & Gaeth, 1988; Seta et al., 2010), our experimental paradigm varied the attribute frame independently from the quantifier magnitude. This allowed us to directly compare effects across quantifier magnitudes. Participants' judgements of small and large amounts of calories were healthier and less healthy respectively relative to energy —and this was supported by the pattern of affect and recommendations inferred about the entire phrase. For instance, framing energy in its negative form, calories, resulted in lower affect and judgement when the amounts were large, but higher affect and judgement when the amounts were small. Quantifier magnitudes thus shape people's perception of an attribute frame, and bears implications for framing theory and research.

The affective encoding account posits that a frame's valence produces a positive or negative affect that influences people's judgement (Levin & Gaeth, 1988). Attribute framing studies assume the magnitude of the quantifier to be irrelevant: 20% fat is negative, as is 5% fat (Kim et al., 2014; Kreiner & Gamliel, 2017). Yet we found that small quantities (in this case, around 20%) changed the valence of the attribute and therefore reversed the subsequent framing effect produced. Further, the main driver of this change appeared to be the shift in affect and judgement of 'calories' rather than 'energy'. Different quantities of calories had clearly opposite valence, but judgements of the quantities of energy remained more stable. The magnitude of the quantifier therefore had different effects on the two frames.

The modification of frame valence by the magnitude of the quantifier could indicate that people are not simply making a quick, affective judgement, but scaling their judgements accordingly. This seems more in line with a pragmatic account of processing the frame, although it does not rule out the possibility that people could have an instinctively affective reaction to an entire quantified phrase, especially if they have been consistently exposed to it (e.g., 'low fat' food labels; Wansink & Chandon, 2006). This might also explain why there was a greater shift in judgement for different quantities of the negative frame of 'calories' than the positive frame of 'energy': people are more familiar with the term in popular media. Nonetheless, the effect of quantifier magnitude suggests at the very least that the affective encoding account would need to be refined to account for the valence indicated by a quantifier.

The moderating role of quantifier magnitude also highlights a methodological implication: the need to control for the magnitude of a quantifier in studying the framing effect. Previous work on attribute framing typically compared small quantities of a negative frame against large quantities of a positive one (e.g., Gamliel & Kreiner, 2013; Kim et al., 2014; Sanford et al., 2002). The resultant effects may not have accounted for how a small quantity could modify the affective nature of the overall phrase. This may explain why certain combinations of attribute frames (e.g., 5% fat vs. 95% lean) displayed smaller, or non-significant effects compared to classic combinations (i.e., 25% fat vs. 75% lean; Kim et al., 2014; Sanford et al., 2002). It could be important to identify a frame's valence at phrase or sentence level, rather than attribute level, to account for how quantifier magnitude can modify valence. Our use of a minimal framing paradigm provides a potential solution to generating framing conditions where quantifier magnitudes can be kept constant. One limitation is that within the energy-calorie frame pair, the negative 'calorie' frame may have been more relevant to a healthiness judgement than the positive 'energy' frame, thus resulting in less change across different quantities of energy than calories. Future work could extend this paradigm to other framing contexts that keep the attribute constant while using synonymous frames with opposite valence. For example, how would one rate the extent of damage to a natural resource that has been 90% 'exploited' vs. 90% 'utilised', and would this be different if it were 10%, or '*slightly*', exploited/utilised?

7.7.4 Practical implications

Our findings also have implications on the communication of food information. The fact that people perceive the same quantity of energy and calories to be different in healthiness indicates that consumers can be led astray by savvy advertising that market products as either 'low in calories' or 'high in energy', but not the other way around. Currently, some legislative guidelines are in place for when a 'low energy' claim may be made, but this does not account for the interchangeable use of this word with 'calories' (Council of the European Union, 2011). In addition, while energy content is mandated on all front-of-pack nutrition labelling (Council of the European Union, 2006; UK Department of Health, 2016), this can be signposted using 'energy' or 'calories' (although 'energy' is the recommended term; UK Department of Health, 2016). To minimise the risk of consumers drawing erroneous conclusions about product healthiness based on the choice of attribute, it may be worth updating policy guidelines to standardise energy terminology. Our work also found that judgements of healthiness based on different quantities of ‘energy’ did not change as much as judgements of healthiness based on different quantities of ‘calories’, which could indicate that labels about ‘energy’ are less useful as a label. In line with tackling obesity, it may also be worth nudging consumers to reduce consumption by requiring the use of ‘calories’ with large quantities to implicitly signal this recommendation.

7.7.5 Conclusion

Over four experiments, we showed that attribute framing is not consistently moderated by format, but it is explained by affect and inferred recommendations about the frame as a whole: attribute *and* quantifier. It is clear that neither an affective encoding nor a pragmatic inference explanation can fully account for the framing phenomenon. Rather, integrating two previously opposed accounts provides a more nuanced view of people’s reasoning processes, and leads us towards a better understanding of the mechanisms behind attribute framing.

Chapter 8: General Discussion

The goal of this thesis was to investigate how a verbal or numerical format affects people's psychological processing of quantifiers. Specifically, I investigated three key areas: (1) Interpretation of quantifiers; (2) Attention to quantified information, and; (3) Evaluation of quantified information. I used two theoretical frameworks, dual-process theory and pragmatic theory, to inform my hypotheses. Across 14 studies, I used a range of methodological and analytical strategies to test the effect of quantifier format. Using online surveys, I investigated how people interpret quantifiers and whether this corresponded with official guidelines. Using decision tasks and eye-tracking methods, I investigated what aspects of information people attend to and how these processes influence decisions. Finally, using vignette-based judgement tasks and mediation analyses, I investigated drivers of people's evaluation of verbal and numerical quantities of food.

This chapter provides a discussion of the overall research in this thesis. I start with an overview of the key findings from chapters 2-7, followed by a brief discussion of how these results inform our understanding of different types of verbal and numerical quantifiers. I then compare the two theoretical frameworks that motivated the studies and argue that the empirical evidence suggests an integrated approach to understanding how people process verbal and numerical quantifiers. From this, I suggest some future directions for research. Finally, I discuss the practical applications of this work. I conclude the chapter with a review of the open science contribution of this thesis, including a reflection on developing open science practices in psychological research.

8.1 Overview of Findings

The first empirical chapter in this thesis investigated how people translate between verbal and numerical quantifiers in the context of nutrition communication (specifically, Guideline Daily Amounts; 'GDAs'). In two studies aimed at replicating past work in a new context, participants translated verbal quantifiers into numerical values with great inter-individual variability, which corroborates previous work done with verbal probabilities (Budescu et al., 2014; Budescu & Wallsten, 1985). Interestingly, the studies also showed that numerical quantifiers were not interpreted as precisely as previous literature suggests (Budescu & Wallsten, 1995). There was substantial variability in perceptions of numerical quantifiers as well as verbal ones. Further, people consistently overestimated verbal quantities of nutrients compared to standard guidelines. Estimation magnitude was not predicted by individual difference measures. However, evidence from Study 2 in this chapter suggested that the valence of a nutrient influenced the magnitude of people's estimations for verbal quantifiers. People estimated quantities of positive nutrients to be greater than the same quantities of negative nutrients.

Chapters 3 and 4 investigated the hypotheses of dual-process theory as applied for the first time to comparing verbal and numerical quantifier processing. Specifically, these four studies tested the postulate that verbal quantifiers elicit a more intuitive processing style than numerical quantifiers (Windschitl & Wells, 1996). There was mixed evidence from four measures of processing style: response time (studies 3-6, total $N = 733$), decision performance (studies 3-6), influence of contextual information (studies 3-6), and performance under concurrent cognitive load (studies 5-6, total $N = 486$). In line with my hypotheses, participants displayed better overall performance for numerical quantifiers than verbal ones (consistent trend across all studies, significant in studies 4-6), and their decisions with numerical quantifiers were less influenced by the contextual information presented (consistent trend across all studies, significant in studies 3, 5, and 6). However, participants did not process verbal quantifiers consistently quicker than numerical ones in the decision task (verbal quantifiers were marginally, and not significantly, slower in studies 3-4, but faster in studies 5-6, however only study 5 found a significant effect). The critical test of whether numerical quantifiers require more analytical processing than verbal ones was that a concurrent memory load should dampen performance of numerical quantifiers, but not of verbal ones (De Neys, 2006). This test (studies 5-6) did not detect a difference between performance at different levels of memory load for either verbal or numerical quantifiers, indicating that both formats did not require much analytical processing in the decision task.

The findings from studies 3-6 showed that a categorical dual-process explanation cannot fully account for the differences in how people make decisions with verbal vs. numerical quantifiers. The results were only partially in line with the predictions of dual-process theory: people made less accurate decisions with verbal quantifiers, which were more influenced by contextual information. While dual-process theory suggests that this is because they are using intuition and displaying intuitive biases, pragmatic theory presents an alternative explanation. In particular, participants could rely on contextual information more for verbal than numerical quantifiers not because they are irrational, but because verbal quantifiers act as a cue to derive more pragmatic inferences from the information. We followed up on this hypothesis in Chapter 5 (study 7, N = 144) by using eye-tracking methodology to trace participants' attention to the contextual information (nutrient) with different quantifier formats. Study 7 found that, in line with the expectations of pragmatic theory and contrary to those of dual-process theory, participants paid more attention to contextual attributes with verbal than numerical quantifiers, and did not give more attention overall to the numerically-quantified information compared to the verbal.

Chapter 6 of this thesis further investigated a pragmatic theory account of differences in verbal and numerical quantifier judgements by revisiting a well-established effect in the JDM literature: attribute framing (Levin et al., 1998). Based on pragmatic theory, which posits that verbal quantifiers serve practical communicative functions beyond conveying the literal meaning of the quantifier (Sanford & Moxey, 2003; Teigen & Brun, 2000), we would expect people to use verbal quantifiers as a pragmatic signal for what information was important to consider. This would result in a stronger attribute framing effect for verbal than numerical quantifiers, and for quantify pairs that provided more ambiguity. The three studies in this chapter (studies 8-10, total N = 1217) had mixed results: verbal quantifiers showed a larger, significant framing effect in only study 8 and only for one of the quantifier pairs (25% fat). However, studies 9-10 (total N

= 882) found that verbal and numerical quantifiers possessed similar directional qualities, which contrasts with previous work that suggested that verbal quantifiers were more directional than numerical ones (Teigen & Brun, 2000). This unexpected and novel finding suggests that numerical quantifiers have pragmatic effects that need to be further investigated. Studies 8-10 also found that participants were sensitive to the magnitude of the quantifiers used in the construction of the attribute frames. Quantifiers with larger magnitudes directed participants' focus to the attribute, whereas quantifiers with smaller magnitudes directed their focus away from it. This supports the postulate of pragmatic theory that people infer more than just the literal meaning communicated —but for both verbal and numerical quantifiers.

Finally, Chapter 7 compared for the first time dual-process and pragmatic theory models of the attribute framing effect for verbal and numerical quantifiers. Across four experiments (studies 11-14, total $N = 1227$), I used mediation analyses to test the predictions of an affective encoding account (derived from dual-process theory) and an inferred recommendation account (derived from pragmatic theory). I found evidence that verbal quantifiers magnified the framing effect slightly, and both people's affective associations with the frame (predicted by the affective encoding account) and their inferences about a speaker's recommendation of the product (predicted by the inferred recommendation account) contributed to how they derived their judgements. When the two accounts were tested in parallel, I found that they were equally strong as mediators of the attribute framing effect for small quantities, but the affective encoding account explained better the attribute framing effect for large quantities.

8.2 Generalisability of Results: Does One Quantifier Compare With Another?

This thesis focused on people's interpretation, attention to, and evaluation of proportional quantifiers, and more specifically, verbal quantifiers denoting amounts that could be expressed in numerical percentages (e.g., 'low' vs. '10%'). This is a small subset of proportional quantifiers, which comprise a broader class of expressions that range in scope (e.g., ‘few’, ‘many’, ‘some’, ‘all’; Sanford et al., 2007). There is in turn a larger family of verbal quantifiers, which include probabilities (e.g., ‘likely’), and frequencies (e.g., ‘common’). Although we took direction from research done in different domains with different quantifiers (probabilities: Collins & Hahn, 2018, frequencies: Knapp et al., 2004; Newstead & Collis, 1987), proportions: Sanford & Moxey, 2003), we note that not all quantifiers are equal in scope —for example, proportional quantifiers require more computation than cardinal quantifiers (e.g., expressions relative to a number, such as ‘less than five’ Szymanik & Zajenkowski, 2009). In this section, I discuss the potential generalisability of my findings to other types of quantifiers often considered in the psycholinguistics literature.

8.2.1 Are all quantifiers equally vague?

A key finding of Chapter 2 was that verbal quantifiers had a vague interpretation: a ‘low %’ can mean a range of numerical values across participants. This finding is consistent with work done with verbal probabilities (Budescu et al., 2009, 2012, 2014; Budescu & Wallsten, 1985) and frequencies (Berry et al., 2002, 2003, 2004; Hamrosi et al., 2012; Knapp et al., 2009a, 2010). When compared to official translations, the tendency to overestimate verbal quantifiers at all levels of the scale (i.e., low was overestimated, and so was high) was more in line with findings from the verbal frequencies literature (e.g., common, very common). The verbal probabilities literature tended instead to find that participants overestimated small probabilities but underestimated large ones compared to an official standard. Though there is less consolidated translation work across the wider range of proportional quantifiers, interpretational vagueness appears to be an enduring characteristic of verbal quantifiers here, too (e.g., Amer & Drake, 2005; Borges & Sawyers, 1974). Therefore, research on format effects for all types of quantifiers should benefit from efforts to maintain equivalence across verbal and numerical quantifiers for each participant. A method suggested in this thesis is to use individual translations that are later displayed in the relevant judgement or decision-making task (see Study 4, Chapter 3; Study 6, Chapter 4; Studies 8-10, Chapter 6; and Studies 12-14, Chapter 7); this method should apply across quantifier types.

8.2.2 Do all verbal quantifiers increase attention to attributes?

Another key finding from this work was that verbal quantifiers increased attention to the attributes they scope: Studies 3-6 (Chapters 2-3) found that participants used this attribute information more in their decision-making, and Study 7 (Chapter 4) suggested that this was explained by the greater attention paid to attributes with verbal than numerical quantifiers. Previous work on gambles —involving probabilities —suggests that people might pay more attention to uncertain events when given verbal probabilities (González-Vallejo et al., 1994), however this has yet to be tested in cases where the event outcome is not also numerical. It is less certain whether similar effects would be found with frequencies. Studies on perceptions of risk frequencies have found that people thought side effects of medication were more risky to health and were less likely to take the medication with verbal risk frequencies than numerical (Berry et al., 2003; Knapp et al., 2004), which could be because this information was more salient with the verbal quantifier. However, this needs to be formally tested.

8.2.3 Would other numerical quantifiers be less susceptible to framing?

Finally, this thesis looked at how quantifiers interact with frames to affect judgement. I proposed that the focusing properties of quantifiers, in particular, verbal quantifiers, contribute to the framing effect. However, the attribute framing effect was only slightly greater with verbal quantifiers, and not consistently so (Studies 8-14, Chapters 6-7). Strikingly, the directional focus of numerical quantifiers matched that of the verbal quantifiers (Studies 9-10, Chapter 6). This is a different finding from similar work for verbal and numerical probabilities (Teigen & Brun, 1995, 2000). One reason could be the difference in methodology: the previous work compared verbal probabilities with their average numerical meanings; my studies compared individually-translated verbal and numerical quantifiers. However, it should also be noted that verbal probabilities have an inherent 'frame' (e.g., 'unlikely' is more obviously negative than 'low'). This may mean that for the same numerical percentage, a verbal probability might offer a different directionality from the proportional quantifier. Verbal frequencies are likely to display this same inherent framing property (e.g., 'uncommon' vs. 'common'), but most proportional quantifiers are unlikely to have an inherent frame: 'few' has negative focus as defined by its linguistic properties (Horn, 1989; Moxey et al., 2001), however one would not readily assume it to be negative in affect as compared to 'a few'. As such, it is possible that the format of probabilities and frequencies could moderate attribute framing effects more than what was found in this thesis for proportional quantifiers.

8.3 A Comparison of Dual-Process and Pragmatic Theory: An Argument for Theory Integration

The empirical work presented in this thesis suggests there is some overlap in how people interpret, attend to, and evaluate verbal and numerical quantifiers. Nonetheless, these processes are also different between formats. In terms of interpretation, people vary in their perceptions of the magnitude of verbal quantifiers, however they also display some variability in perceiving numerical quantifiers (Study 1, Chapter 2). People do not process one format significantly quicker than the other (Studies 3-6, Chapters 3-4), nor does a concurrent cognitive load task interfere more with one format than the other (Studies 5-6, Chapter 4). In terms of information use, people are more influenced by contextual information with verbal quantifiers, particularly when this information supports their existing assumptions (Studies 3-6, Chapters 3-4). Furthermore, the greater influence of contextual information appears to be driven by a tendency to pay more attention to contextual attributes with verbal than numerical quantifiers (Study 7, Chapter 5). The properties of these attributes, for example people's positive or negative associations with them, then influence people's judgements of the quantified information (Studies 8-14, Chapter 6-7).

These results do not fit easily within a single theory. In fact, a recurring theme throughout this thesis has been that each theory, dual-process or pragmatic, explains part of a study's findings, by each focusing on a different level of the judgement or decision process. Dual-process theory addresses the nature of information processing. It posits that people can be more or less intuitive (or analytical) in their judgement and decision-making, with each style associated with certain behavioural correlates (e.g., intuitive processing is automatic and quick, but also more prone to intuitive bias Evans, 2008; Kahneman, 2011). Explanations deriving from dual-process theory, such as the affective encoding account (Levin & Gaeth, 1988), suggest an automatic priming process whereby the level of positive (or negative) affect associated with an attribute predisposes a person to judge the quantified information positively (or negatively). If verbal quantifiers are more intuitively processed than numerical ones, dual-process theory could account for the greater influence of contextual information on decisions with verbal quantifiers, as observed in Studies 3-6 (Chapters 3-4). Dual-process theory could also explain some of the variation in attribute framing effect sizes between verbal and numerical formats, as observed in Studies 8-14 (although this moderated effect was not consistently replicated across studies; Chapters 6-7). However, dual-process theory is not the only theory that could explain these results. The greater influence of contextual information, in both decision tasks and framing situations, could be a feature of verbal quantifiers, which are easier than numerical quantifiers to integrate with the context (Moxey & Sanford, 1993). Further, dual-process theory also does not account for why people gave attributes and quantifiers more visual attention for verbal than numerical quantifiers in Study 7 (Chapter 5), or why framing effect sizes differed across quantifiers of different magnitudes in Studies 8-14 (Chapters 6-7).

Pragmatic theory addresses the purpose of information processing. It posits that language serves a practical function beyond simple informational content (Kess, 1992). The pragmatic inference account, which is derived from pragmatic theory, suggests that people make rational inferences about quantities (which could be automatic or deliberate). According to this account, people infer that certain expressions were chosen to communicate additional information about the quantity (Sher & McKenzie, 2006), such as the speaker's recommendation for what to choose (van Buiten & Keren, 2009). Based on the postulate that verbal quantifiers have greater pragmatic signalling, pragmatic theory can account for why people paid more attention to the attributes with verbal quantifiers in Study 7 (Chapter 5). It can also explain why people's judgements and decisions were sensitive to changes in both the attributes (i.e., the frame) and quantifiers (i.e., the format and magnitude) in Studies 8-14 (Chapter 6-7). However, a key empirical prediction was that the attribute framing effect would be consistently magnified with verbal compared to numerical quantifiers. This prediction, which should follow from pragmatic theory, was not supported. This calls into question the assumption from the theory that verbal quantifiers have a larger pragmatic signal than numerical ones. Indeed, Studies 9-10 (Chapter 6) found that numerical quantifiers possessed comparable pragmatic signals to verbal ones in terms of directional focus. This challenges the view that numerical quantifiers are more ambiguously focused than verbal ones (Teigen & Brun, 2000), at least for certain quantities (specifically 5% in Studies 8-10).

In the final study of this thesis (Study 14, Chapter 7), a parallel analysis of explanations derived from both theories —dual-process and pragmatic —showed that both affect and inferred recommendations contributed to variation in participants' judgements. The explanatory strength of the affective or pragmatic account varied across the magnitude and format of the quantifier. Notably, the results also suggested that elements of the two theoretical accounts should be refined such that both accounts could offer stronger explanations for the data. The affective encoding account needs to consider that quantifiers can modify the affect generated by a frame. The pragmatic inference account needs to consider that numerical quantifiers in certain contexts may possess unambiguous directionality. Altogether, the findings from this thesis suggest that an integrated theoretical explanation that considers both dual-process and pragmatic theories is most useful in furthering our knowledge of the processes at play when people encounter quantified information. In the next sections, I discuss further how further research could be undertaken to improve and integrate dual-process and pragmatic theories for understanding quantifiers.

8.4 Future Directions

This thesis offers empirical and methodological advances to the study of quantifier processing in judgement and decision-making. Over a set of 14 experiments, it provides a systematic test of the differences in processing verbal vs. numerical quantifiers. I confirmed that numerical translations of verbal quantifiers used in nutrition communication exhibit similar variability to previous translation work done with verbal probabilities, and provide data on these translation values (Chapter 2). This data can be subsequently used as a baseline for future research. I showed that people exhibit differences in attention (Chapter 5), information use, and decision biases (Chapters 3-4) for verbal and numerical quantifiers. Finally, I provided evidence that both dual-process and pragmatic theories explain these differences (Chapter 5-7).

Methodologically, I added to previous research by employing novel tasks to test our empirical questions. I extended methods in verbal-numerical quantifier translations by soliciting back-translations (numerical-verbal) and using visual analogue scales to test participants' interpretations of both quantifier formats (Study 1 in Chapter 2). I created a new decision task (Studies 3-6 in Chapters 3-4) that could measure reaction times and decision performance with quantifier calculations in the lab and online; crucially, this task does not rely on numerical presentation of attribute values. I developed a translation method to control for inter-individual variability when conducting comparative experiments between quantifiers formats (Study 4, Chapter 3; Study 6, Chapter 4; Studies 8-10, Chapter 6; and Studies 12-14, Chapter 7). Finally, I tested a minimal framing paradigm to isolate the effects of quantifier magnitude and format in the attribute framing effect (Studies 11-14, Chapter 7).

Building on these empirical and methodological contributions, I offer some suggestions for the further study of quantifiers in judgement and decision-making. In the following subsections, I outline three areas for future research that follow from our work, which consider (1) dual-process theory; (2) pragmatic theory, and; (3) the integration of both theories.

8.4.1 Verbal and numerical quantifiers may be processed along a continuum

The present work suggests that the processing distinction between verbal and numerical quantifiers is not categorical. Building on previous work and postulates about the intuitive and analytical nature of the two formats (Windschitl & Wells, 1996), we used a basic form of traditional dual-process theory (Evans, 2008) as a starting framework for our investigation. However, it is important to acknowledge that dual-process theory has evolved into different forms and models (e.g., default-interventionist: Sloman, 1996; parallel-processing: Glöckner & Hodges, 2010; cognitive continuum: Hammond, 1980), with a plethora of debate over how the concepts of intuition and analysis should be defined and modelled (e.g., Betsch & Glöckner, 2010; Glöckner & Witteman, 2010a; Kahneman & Klein, 2009; Kruglanski & Gigerenzer, 2011; Osman, 2004). One example is the classification of 'heuristics', which to some research programmes is indicative of intuitive processing (e.g., heuristics and biases programme: Kahneman, 2003; heuristic-systematic systems: Zuckerman & Chaiken, 1998). In contrast, other programmes argue that how people use heuristics —automatically or deliberately —determines whether heuristics-based decisions are intuitive or analytical (Gigerenzer & Goldstein, 1996; Glöckner et al., 2014). Another point of contention in this debate is which of the many pairs of characteristics proposed in the dual-process literature is necessary to define an intuitive or analytical process (Evans & Stanovich, 2013). Bago & De Neys (2017) suggest that a defining characteristic of intuition is the ability to operate under concurrent interference with working memory, though Bago & De Neys (2019) caution that it is still difficult to differentiate intuitive and analytical decision outcomes using this method. Others advocate the use of multiple measures to capture different aspects of the two styles (Horstmann et al., 2010).

Based on the methodological suggestions from the literature, we used a range of measures to test for intuitive processing. Our empirical evidence showed that compared to numerical quantifiers, the processing of verbal quantifiers relies more on contextual information and in general leads to more errors when this context is counterintuitive, such as identifying the overconsumption of minerals. However, people do not process verbal quantifiers consistently quicker, with less effort, or crucially, with less interference from additional cognitive load. These findings do not fit well with the traditional 'intuitive-or analytical' dual-process framework (Evans, 2008; Kahneman, 2011). However, the other varieties of dual-processing theory outlined above may better inform our understanding of how intuition and analysis are involved in processing quantifiers. An alternative approach that has run in parallel to the popularised two-systems programme (Kahneman, 2011; Sloman, 1996) is to view processing on a continuum, where different levels of intuition and analysis contribute to a final judgement or decision (Hammond, 1980, 1981). This is more so the case when we consider that decision-making using quantifiers may comprise of multiple processing steps, including attending to the information, deriving meaning from it, judging the information, integrating it with other information in the task or from memory, and finally reaching a decision (Shaw, 1982). My research focused on several of these processes (attention: Chapter 5; meaning: Chapters 2, 6, and 7; judgement: Chapters 5-7; decisions: Chapters 3-4). The mixed findings about the extent of intuitive or analytical processing that applies to each quantifier format may reflect a cognitive versatility for switching processing styles across different steps of a decision task. A perspective of decision processes that takes into account multiple steps may provide a middle ground between the categorical and continuous views of dual-process theory. If each processing step is either intuitive or analytical (as stipulated by the categorical model), the final observed outcome may still reflect a continuum of intuitive and analytical processes (as stipulated by the continuous model) depending on how many of the interim steps were intuitive and how many were analytical. A future direction for this work would therefore be to investigate specifically the components of a decision task with respect to how quantifier format differentially affects each processing step. For example, varying a memory load at different points of information presentation in a decision task could identify which step of the task the load affects.

8.4.2 People can derive multiple pragmatic inferences from verbal and numerical quantifiers

The present work identified that people derive pragmatic inferences from both verbal and numerical quantifiers. Specifically, we identified two types of pragmatic information that contribute to people's judgements of quantified information: directional focus (Chapter 6) and implicit recommendations (Chapter 7). However, the pragmatic theory literature suggests that there are more types of inferences that people could derive from a quantified phrase. For example, people can infer that the phrase, 'a two-thirds of the people will die', means at least two-thirds (i.e., two-thirds or more; Mandel, 2015). Alternatively, people can infer the reference point by which to compare the quantity, whether relative to a previous trend or expectation (e.g., a glass half-full is inferred to be previously empty; Sher & McKenzie, 2006), or to a comparative option (e.g., 70% chance of flood means the flood is more likely than the landslide; Windschitl et al., 2017).

I have not been able to explore the full range of possible pragmatic inferences people could make within this thesis. However, I anticipate that the quantifier format could cue different inferences in each category. For example, it could indicate a different range of estimates that the quantity should fall into: one could infer a verbal quantifier to indicate the mid-point of a range (Budescu & Wallsten, 1995), but a numerical quantifier to indicate a lower-bound of a range (Mandel, 2015). Paired with a frame, the quantifier format could highlight the reference point of the speaker, such that one could infer that a positively-framed verbal quantifier such as 'high rate of success' suggests an improved success rate more than '70% success'. Finally, the quantifier format could make clearer comparisons to other options: one could infer that a verbal quantifier such as 'high fat' (as opposed to '75% fat') suggests higher fat than other foods. Thus, a future direction for this work lies in systematically addressing the different types of pragmatic inferences that occur when encountering quantified information (e.g., range boundaries, comparisons to trends or alternative options), and comparing the applicability of these inferences to different quantifier formats.

8.4.3 Delving further into the interplay of dual-process and pragmatic theories

My research suggests that both dual-process and pragmatic theories have an explanatory function in quantifier processing. Across different studies, I was able to support hypotheses drawn from both theories, with the strength of each explanation varying across studies.

Studies 3-6, assessing the use of contextual information in decision-making, found greater use of the context for verbal than numerical quantifiers (see Chapters 3 and 4). Study 7, addressing attention to aspects of quantified information, found greater attention on context (nutrients) with verbal than numerical quantifiers (see Chapter 5). In these chapters, I interpreted the results as a function of people's inability to suppress an intuitive response with verbal quantifiers —a dual-process explanation. The use of contextual information affected performance with verbal more than numerical quantifiers, which fits the traditional dual-process conception of intuitive biases (Kahneman, 2011). My decision task presented some contextual elements that suited the nature of the task (e.g., it is typical to decide if one has exceeded a fat target) and others that were more counter-intuitive (e.g., it is more typical to decide if one has achieved a minerals target rather than exceeded one). Both conditions allowed for intuitive responses (minerals –healthy; fat –unhealthy) and counter-intuitive ones (minerals –unhealthy; fat –healthy) within the task context. Therefore correct responding in the counter-intuitive conditions (more so for minerals than fat) depended on being able to decouple the quantifier from the context, consistent with analytical processing. However, one must also consider these results through the lens of pragmatic theory. A key tenet of pragmatic theory is that verbal quantifiers scope items in relation to their context (Moxey & Sanford, 1993; Weber & Hilton, 1990). Although I accounted for individuals' interpretations of verbal quantifiers relative to fat and minerals, the task itself (deciding if a target had been exceeded) would be additionally counter-intuitive for the verbal quantities of minerals. If verbal quantifiers perform an additional pragmatic function of scoping the nutrient relative to its natural decision type (i.e., 'high % minerals' is easily understood in relation to helping achieve a target), one would reasonably falter because the task context is less relevant to the typical usage of the verbal quantifier. A future direction for research might therefore be to vary the task (e.g., deciding whether one has adequately reached a target instead of whether one has exceeded it) to ascertain whether the opposite pattern occurs (i.e., more counter-intuitive errors with fat than minerals on top of more counter-intuitive errors with verbal than numerical quantifiers). This could tap into the possible pragmatic processes at play.

The final empirical study in this thesis sought to integrate and compare both dual-process and pragmatic theories in a model of attribute framing. Study 14, addressing the role of affect or inference in judgements, found comparable evidence for both accounts with low quantities, but stronger evidence for the dual-process account with high quantities (see Chapter 7). One limitation in this study, however, was that the mediators we tested (attention, affect, and inferences) were not independently manipulated in the experimental design, thus I could not provide a confirmatory test of their role. I see future potential to further this work by specifically manipulating a hypothesised causal mediator in order to experimentally test its effect. For example, one could present people with an imaginary attribute (e.g., 'G' vs. 'non-G'). If the framing effect is produced only if this attribute is previously paired with positive or negative stimuli, one could then trace the effect to the affect associated with that attribute.

A further issue that has yet to be addressed is how dual-process and pragmatic theories could feed into one another. While I compared both theories and their contribution to understanding quantifier processing, the explanations need not be mutually exclusive. Pragmatic inferences, for example, could proceed intuitively or analytically. Thus far, the literature is inconclusive about whether language inferences are effortful (Bott & Noveck, 2004; De Neys & Schaeken, 2007; Dieussaert et al., 2011; Zhao et al., 2015). Further, this research on the automaticity of inferences is concentrate on scalar inferences (i.e., practical conclusions about the use of quantifiers like 'some' and 'all'), so it is also debatable whether other types of quantifier inferences, such as those highlighted in section 8.4.2, require effort. We therefore see as an important future direction for our work further investigation into the interplay of pragmatic and cognitive factors in quantifier processing. For example, how might affective associations influence the inferences people draw from a quantified phrase? This investigative direction has the potential to inform our theoretical and practical knowledge of how language and cognition interact in judgement and decision-making.

8.5 Practical Applications

Although this thesis has a primary focus on theory and cognitive processing, its findings bear several practical applications, particularly in the applied context of nutrition labelling. In this section, I summarise three practical challenges raised by our results. These challenges address (1) whether consumers' interpretations of nutrition label give them accurate nutrition information; (2) whether consumers' focus of attention on different aspects of nutrition labels leads them towards healthful decisions, and; (3) whether consumers can infer helpful additional information from a nutrition label.

8.5.1 Consumer interpretations of nutrition labels may be misleading

My work shows that verbal quantities of nutrients on nutrition labels may lead to overestimations of the numerical GDA they refer to. Studies 1 and 2 (Chapter 2) showed that although participants consistently ranked verbal quantifiers of different magnitudes in the same order, individual interpretations of each quantifier varied greatly across participants and, more crucially, fell well above stipulated ranges in standard guidelines. While it is true that verbal banding is not always determined by the GDA percentage (e.g., for fat, it is calculated by grams of fat per 100g of product; UK Department of Health, 2016), GDAs are intended to be the more understandable numerical format that consumers are expected to rely on (Grunert et al., 2010b; Rayner et al., 2004). When the two are presented in conjunction (see Figure 8.1) consumers are likely to interpret verbal quantifiers in terms of a percentage GDA contribution. This is even more likely because numerical GDA values are standardised and thus comparable across nutrients. Rather than remembering that low fat is 3g per 100g, but low sugar is 5g per 100g, it is quicker and simpler to associate 'low' with a GDA percentage of, say, 5-10%. Further, for certain nutrients (e.g., fibre, protein), technical guidance does calculate verbal quantifiers based on how much the product contributes to one's GDA (UK Department of Health, 2016). It is therefore important to consider what consumer perceptions are in terms of GDAs.

The fact that I was unable to predict variation in magnitude estima-

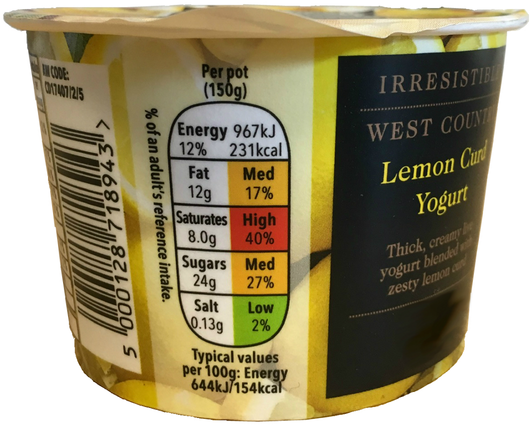

Figure 8.1. Use of verbal quantifiers alongside numerical GDA % quantifiers in a real-world product. Photograph taken and edited by the author to remove brand information.

tions for verbal quantifiers suggests that general interpretations of quantifiers are ingrained and not improved by familiarity with labels or individual attitudes towards healthy eating. More evidence is necessary to ascertain if targeted education about nutrition labelling can indeed adjust people's interpretations. However, findings from other domains show that even when people are given verbal quantifiers with their stipulated numerical ranges side-by-side, their translations of the verbal quantifiers are still misaligned with the official standard (Budescu et al., 2012). Further, the need to pre-educate people about nutrition labelling contradicts the goal of interpretive nutrition labelling, which is to provide simple, clear, and easily understandable information (Malam et al., 2009). A challenge for nutrition labelling is therefore to consider whether verbal banding for nutrient quantities should take into account what natural interpretations people would make about the verbal quantifiers.

8.5.2 Quantifier format changes people's attention to different information on a nutrition label

We found that verbal quantifiers increased the amount of attention the nutrient on a label received. Verbal quantifiers could thus magnify or reduce the perception of a food's healthiness compared to numerical ones, depending on which nutrient it is paired with. Studies 3-6 (Chapters 3-4) suggest that decisions with verbal quantifiers are influenced by assumptions that a nutrient is healthy or unhealthy, more than decisions with numerical quantifiers. Study 7 (Chapter 5) showed that participants looked longer at a nutrient when the quantifier was verbal than when it was numerical, which could explain why people relied more on the nutrients to make decisions with verbal quantifiers: they paid more attention to it. This difference in attention focus has deeper implications for nutrition labelling.

Studies 3-6 (Chapters 3-4) highlighted two types of mistake that people can make when deciding if it is healthy to eat something. People can mistake healthy amounts for unhealthy or unhealthy amounts for healthy. In the GDA decision task, participants made the most errors when mistaking unhealthy amounts of minerals as healthy. Paying too much attention to positive nutrients can thus have detrimental consequences in practice. People can overestimate a food's healthiness (Ebneter et al., 2013; Gravel et al., 2012) and consume too much of it (Wansink & Chandon, 2006). By focusing attention on positive nutrients, verbal labels could reinforce a 'health halo effect', where positive food attributes cause the entire food to be judged as healthy irrespective of its actual nutritional value (Roe et al., 1999). This problem is illustrated in Figure 8.2, where the verbal quantifier attached to the 'high fibre' claim may obscure the numerical percentage of sugar included.

From a public health perspective, whether increased attention to a nutrient is beneficial depends on the message, or goal, one wishes to put forth. If the message is that people need to eat less (Crockett et al., 2018; Storcksdieck genannt Bonsmann & Wills, 2012), the potential for verbal quantifiers to reinforce the health halo effect could be detrimental for making healthy choices. However, if the goal is to promote consumption of beneficial nutrients (e.g., fibre; Guiné, R. P. F., Duarte, J., Ferreira, M., Correia, P., Leal, M., Rumbak, I., ... Straumite, E., 2016), focusing on attribute information may at least prompt people to include more of these beneficial nutrients in their diet. The challenge for nutrition labelling is thus to consider how attention patterns for different quantifiers can be harnessed to direct people to information that is most predictive of healthful decisions.

Figure 8.2. Use of verbal and numerical quantifiers for different nutrients on a real-world product. Photographs were taken and edited by the author to remove brand information.

8.5.3 Framing and format provide implicit cues about how consumer should judge food healthiness

Studies 8-10 (Chapter 6) showed that people rely on these cues most when there is ambiguity surrounding the quantities they are trying to judge. If one does not know whether a 25% fat beef is healthy, one could infer from the use of the attribute 'fat' that the communicator views the product less healthily than if they had described it as '75% lean'. Chapter 7 further investigated what people inferred about a communicator's recommendations, and found that participants regarded products described as 'low calorie' and 'high energy' to be recommended to them by the speaker more than when the same quantities were 'low energy' and 'high calorie'. This corresponded with their perceptions of the product's healthiness (Studies 13-14, Chapter 7). This sort of linguistic framing is commonly found in food and nutrition labelling (see Figure 8.3 for an example, where the Front-of-Pack label on one snack —rice cakes —is described in terms of its calorie content, while the other —biscuits —is described in terms of its energy content). These variations in information frames could present a very different view of the healthiness of two products despite the calorific content being similar. It is certainly possible to mention both frames in communication, which is being advocated, for example, in medical and health communication (Gigerenzer & Kolpatzik, 2017). However, this would require more space on a food label, increase the complexity of information (Roberto & Khandpur, 2014), and may even appear unduly informative (Grice, 1975). In this context, a truly objective method of presenting information may not be practicable. Practitioners and policy-makers might thus need to consider how subtle linguistic variations in format and frame can best serve population health objectives and use this to inform the development of more effective nutrition labelling guidelines.

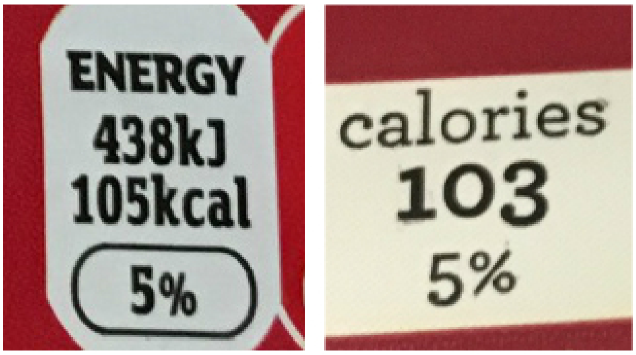

Figure 8.3. Example of energy and calorie frames on Front-of-Pack nutrition labels in the UK. Photographs were taken and edited by the author to remove brand information.

8.6 A Reflection on Open Science for Empirical Research in this Thesis

The methods of this thesis were motivated by the need to contribute to good scientific practice in psychological research. To that end, I considered recent recommendations from open science initiatives in conducting my studies (Open Science Collaboration, 2014). Specifically, Studies 3-14 were all pre-registered with their hypotheses, research design, and analysis plan prior to conducting the research. I conducted power analyses to determine most of the sample sizes and ensure the respective studies were appropriately powered to detect anticipated effect. Where possible the empirical findings from previous studies provided input for expected hypotheses and effect sizes in subsequent studies. Finally, I used the

Open Science Framework to make the pre-registrations, study materials, data, and analysis syntax openly available. These practices promote reproducibility of research, however they presented a number of challenges over the course of the thesis. As processes in open science are evolving and the open science movement continues to spread, it is important to reflect on how recommendations can be refined and presented as advice to fellow researchers (Laws, 2016; Munafò et al., 2017). As such, this section discusses the challenges encountered in performing each of the recommended open science best practices.

8.6.1 Challenges in pre-registration

The barriers to pre-registration faced in the first empirical chapter in this thesis were a misconception about whether to use pre-registration in survey data, and constraints on time and participant availability. The first two studies in this thesis were not pre-registered because their main purpose was to establish means and standard deviations of participants' interpretations of the food quantifiers, which would be crucial to the design of subsequent experiments. I mistakenly assumed that pre-registration was not relevant as the primary purpose was not testing a classic experimental manipulation.

Similar misconceptions about whether pre-registration is applicable to a study are not uncommon (Nosek et al., 2018). In order to correct these misconceptions, it is important therefore to stress that hypotheses need not take a specific form. In the case of my studies in Chapter 2, understanding that the rationale for collecting the measures in the surveys could constitute my hypotheses would have helped in crafting a pre-registration. Improving the versatility of pre-registration protocols could also help in making pre-registration a more efficient process. For example, the simplest pre-registration site (AsPredicted.org) has a simple eight-question procedure, however co-author approval processes only allow approval or rejection, and not editing by co-authors, which means that if one author rejects the submission, the pre-registration must be completed all over again. Researchers on collaborative work may thus be wary of the time cost of the procedure, especially since it is not certain at the beginning whether a pre-registration is time-stamped to its creation, or its final approval, and whether all authors in a final paper submission must be listed in the pre-registration. Definitive advice and recommendations on these fronts could be useful in promoting pre-registration practices more widely.

A further challenge in the pre-registration process was how to account for multiple studies following up on a preliminary hypothesis. This was the case, for example, for Chapters 3, 4, 6, and 7. As each experiment built on the previous, I pre-registered each one in turn. However, this ultimately meant that meta-analysing results across all experiments once data has already been conducted could appear as post-hoc compared to each pre-registered study. One could plan an overall package of studies, however in such a case, it would be difficult to anticipate unexpected findings in earlier studies that suggest design changes in subsequent studies. Nonetheless, pre-registration still provides an overall benefit of maintaining transparency across each step of a study programme. However, as the open science movement grows, it may be beneficial to distinguish study-level and programme-level pre-registrations to accommodate different types of research.

8.6.2 Challenges in power analysis

Conducting power analyses to set a target sample ahead of data collection has the advantage of ensuring the sample is large enough to detect the effect of interest (Button et al., 2013). In theory, this seems a simple process of conducting the analysis and collecting data for exactly the number of participants indicated. In practice, this is more complicated. First, determining the effect size can be a challenge, especially if the effect in question does not yet have reported empirical data to extrapolate from. In such cases, I found a conservative approach to be useful. By assuming effects to be small, even for those with previously reported effect sizes (since publication bias can exaggerate effect sizes; Rosenthal, 1979), one has a better chance of achieving sufficient power.

Second, even if one has an expectation for the overall effect size, the power analysis can be highly complex if advanced statistical analyses (e.g., multi-level modelling) are planned (Scherbaum & Ferreter, 2008). Often the researcher must make a large number of predictions about parameters for each anticipated fixed and random effect, which is difficult to ascertain prior to conducting the study. In such cases, a conservative approach by estimating the sample size required for a proxy analysis (e.g., of aggregated vs. clustered data) can provide a reasonably close estimate. Small-scale pilots can also be useful to establish potential parameters.

Third, data collection is often constrained by external factors such as time pressures, and the cost of participants' time. An eye-tracking study, for instance, is more costly and has limited access to participants compared to an online survey. Checking feasibility of the methods with required sample numbers is thus an important, but often neglected aspect of designing a study.

8.6.3 Challenges in sharing study materials

The growth of open access archives such as the Open Science Framework is a great boost to sharing materials and data. The Center for Open Science recommends making available research materials, data, analytical code, and preregistrations (Center For Open Science, 2018). Data come in different format depending on the nature of the study. For instance, the studies in this thesis were initially shared in their original file formats on the OSF (e.g., as I worked primarily with SPSS for analysis, the data would be provided in this file format). A major problem was that SPSS is not a free programme, thus this format would not be suitable for analysis to researchers without institutional access to these programmes. I have subsequently converted files to comma-separated-value formats, which are more accessible. While this task can be embraced as part of the overall research process, it does require time and effort to convert the files, but also create explanatory material of the variables (which are encoded in an SPSS file but not a comma-separated-value file).

A solution proposed by proponents of open science is to use open source software such as R, which offer the ability to share data and analyses from start to finish. However, this requires a substantial knowledge of coding that may be daunting to researchers, especially those performing qualitative research. While programming skills are greatly beneficial for psychological research, they should not be a barrier to practising open science. It would thus benefit the open science agenda to suggest data sharing protocols that can be flexible across disciplines.

Further, providing clear examples of expectations, such as how and where to provide study information, could therefore improve the prevalence of data sharing among researchers.

8.7 Conclusion

Quantified information, such as the amount of nutrients in food, can be communicated using verbal or numerical quantifiers. The use of these different formats produces different patterns of attention, judgement, and decision-making. These differences can be explained by a combination of cognitive processing styles and pragmatic language inferences. The work presented in this thesis suggests that while people often rely on contextual information (for example, the identity of the nutrient) that is peripheral to the quantitative judgement, especially with verbal quantifiers, they are also capable of extracting implicit information about the quantity from that context (for example, that a communicator recommends a 'high energy' food but not a 'high calorie' one). Future work that explores the step-by-step processing within a decision sequence, the breadth of inferences that could be drawn from quantified communication, and the way cognitive factors interact with linguistic processing could provide more insight into the decision processes involved with quantified information. The insights from this thesis are informative to applied communications, for example in a nutrition context, where quantifier formats could be selected to direct people to informational aspects that predict a healthier choice, and ensure that the communicator's intent is aligned with what people infer about the nutritional quantities.

References

Altmann, G. (1998). Ambiguity in sentence processing. *Trends in Cognitive Science*, 2, 146–152.

Amano, K., Goda, N., Nishida, S., Ejima, Y., Takeda, T., & Ohtani, Y. (2006). Estimation of the timing of human visual perception from magnetoencephalography. *Journal of Neuroscience*, 26, 3981–3991.

Amer, T. & Drake, P. (2005). Qualitative expressions of magnitude: The auditor's responsibility. *American Journal of Business*, 20, 29–36.

Ang, N. & Trotman, K. (2015). The utilization of quantitative and qualitative information in groups' capital investment decisions. *Behavioral Research in Accounting*, 27, 1–24.

Ayal, S., Rusou, Z., Zakay, D., & Hochman, G. (2015). Determinants of judgment and decision making quality: The interplay between information processing style and situational factors. *Frontiers in Psychology*, 6, 1088.

Bago, B. & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, 158, 90–109.

Bago, B. & De Neys, W. (2019). The smart system 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking and Reasoning*.

Berry, D. (2006). Verbal labels can triple perceived risk in clinical trials. *Therapeutic Innovation and Regulatory Science*, 40, 249–258.

Berry, D., Knapp, P., & Raynor, T. (2002). Is 15 per cent very common? informing people about the risks of medication side effects. *The International Journal of Pharmacy Practice*, 10, 149–151.

Berry, D., Raynor, D., & Knapp, P. (2003). Communicating risk of medication side effects: an empirical evaluation of eu recommended terminology. *Psychology, Health and Medicine*, 8, 251–263.

Berry, D., Raynor, T., Knapp, P., & Bersellini, E. (2004). Over the counter medicines and the need for immediate action: A further evaluation of european commission recommended wordings for communicating risk. *Patient Education and Counseling*, 53, 129–134.

Betsch, C. (2004). Preference for intuition and deliberation (pid): An inventory for assessing affect- and cognition-based decision-making. *Zeitschrift fÃŒr Differentielle und Diagnostische Psychologie*, 25, 179–197.

Betsch, T. & Glöckner, A. (2010). Intuition in judgment and decision making: Extensive thinking without effort. *Psychological Inquiry*, 21, 279–294.

Beyth-Marom, R. (1982). How probable is probable? a numerical translation of verbal probability expressions. *Journal of Forecasting*, 1, 257–269.

Białek, M. & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision Making*, 12, 148–167.

Blalock, S., Sage, A., Bitonti, M., Patel, P., Dickinson, R., & Knapp, P. (2016). Communicating information concerning potential medication harms and benefits: What gist do numbers convey? *Patient Education and Counseling*, 99, 1964–1970.

Blitstein, J. & Evans, W. (2006). Use of nutrition facts panels among adults who make household food purchasing decisions. *Journal of Nutrition Education and Behavior*, 38, 360–4.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113, 700–765.

Borges, M. & Sawyers, B. (1974). Common verbal quantifiers: Usage and interpretation. *Journal of Experimental Psychology*, 102, 335–338.

Borgmeier, I. & Westenhoefer, J. (2009). Impact of different food label formats on healthiness evaluation and food choice of consumers: A randomized-controlled study. *BMC Public Health*, 9, 184.

Bott, L. & Noveck, I. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, 51, 437–457.

Brun, W. & Teigen, K. (1988). Verbal probabilities: Ambiguous, context-dependent, or both? *Organizational Behavior and Human Decision Processes*, 41, 390–404.

Budescu, D., Broomell, S., & Por, H.-H. (2009). Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychological Science*, 20, 299–308.

Budescu, D., Karelitz, T., & Wallsten, T. (2003). Predicting the directionality of probability words from their membership functions. *Journal of Behavioral Decision Making*, 16, 159–180.

Budescu, D., Por, H.-H., & Broomell, S. (2012). Effective communication of uncertainty in the ipcc reports. *Climatic Change*, 113, 181–200.

Budescu, D., Por, H.-H., Broomell, S., & Smithson, M. (2014). The interpretation of ipcc probabilistic statements around the world. *Nature Climate Change*, 4, 508–512.

Budescu, D. & Wallsten, T. (1985). Consistency in interpretation of probabilistic phrases. *Organ Behav Hum Decis Process*, 36, 391–405.

Budescu, D. & Wallsten, T. (1990). Dyadic decisions with numerical and verbal probabilities. *Organizational Behavior and Human Decision Processes*, 48, 240–263.

Budescu, D. & Wallsten, T. (1995). Processing linguistic probabilities: General principles and empirical evidence. *Psychology of Learning and Motivation*, 32, 275–318.

Budescu, D., Weinberg, S., & Wallsten, T. (1988). Decisions based on numerically and verbally expressed uncertainties. *Journal of Experimental Human Perception and Performance*, 14, 281–294.

Busemeyer, J. & Johnson, J. (2007). Computational models of decision making. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of decision making* (pp. 133–154). Oxford, UK: Blackwell Publishing Ltd.

Button, K., Ioannidis, J., Mokrysz, C., Nosek, B., Flint, J., Robinson, E., & Munafò, M. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14, 365–376.

Byrd-Bredbenner, C., Alfieri, L., & Kiefer, L. (2000). The nutrition label knowledge and usage behaviours of women in the us. *Nutrition Bulletin*, 25, 315–322.

Campos, S., Doxey, J., & Hammond, D. (2011). Nutrition labels on pre-packaged foods: A systematic review. *Public Health Nutrition*, 14, 1496–506.

Carrigan, N., Raynor, D., & Knapp, P. (2008). Adequacy of patient information on adverse effects: an assessment of patient information leaflets in the uk. *Drug Safety*, 31, 305–312.

Celnik, D., Gillespie, L., & Lean, M. E. J. (2012). Time-scarcity, ready-meals, ill-health and the obesity epidemic. *Trends in Food Science and Technology*, 27, 4–11.

Center For Open Science (2018). Policy and culture change at cos: A strategic overview.
https://cos.io/policy/.

Childers, T. & Viswanathan, M. (2000). Representation of numerical and verbal product information in consumer memory. *Journal of Business Research*, 47, 109–120.

Collins, P. J. & Hahn, U. (2018). Communicating and reasoning with verbal probability expressions. *Psychology of Learning and Motivation*, 69, 67–105.

Council of the European Union (2006). Regulation (ec) no 1924/2006 on nutrition and health claims made on foods. `http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32006R1924`.

Council of the European Union (2011). Regulation (ec) no 1169/2011 on the provision of food information to consumers. `https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32011R1169`.

Cowburn, G. & Stockley, L. (2005). Consumer understanding and use of nutrition labelling: A systematic review. *Public Health Nutrition*, 8, 21–28.

Craig, R. & Mindell, J. (2014). Health survey for england 2013: Health, social care and lifestyles. summary of key findings.

Cramer, A., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R., Waldorp, L., & Wagenmakers, E.-J. (2016). Hidden multiplicity in exploratory multiway anova: Prevalence and remedies. *Psychonomic Bulletin and Review*, 23, 640–647.

Crockett, R., King, S., Marteau, T., Prevost, A., Bignardi, G., Roberts, N., Stubbs, B., Hollands, G., & Jebb, S. (2018). Nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption. *Cochrane Database Systematic Reviews*, 27, CD009315.

Cumming, G. (2013). The new statistics: Why and how. *Psychological Science*, 25, 7–29.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer & P. M. Todd (Eds.), *Simple heuristics that make us smart* (pp. 97–118). New York, NY: Oxford University Press.

De Neys, W. (2006). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *The Quarterly Journal of Experimental Psychology*, 59, 1070–1100.

De Neys, W. (2017a). Bias, conflict, and fast logic: Towards a hybrid dual process future? In W. De Neys (Ed.), *Dual Process Theory 2.0* (pp. 47–65). London, UK: Routledge.

De Neys, W. (2017b). *Dual Process Theory 2.0.* London, UK: Routledge.

De Neys, W. & Schaeken, W. (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology*, 54, 128–133.

Decruyenaere, M., Evers?Kiebooms, G., Denayer, L., & Van den Berghe, H. (1992). Cystic fibrosis: community knowledge and attitudes towards carrier screening and prenatal diagnosis. *Clinical Genetics*, 41, 189–196.

DeStefano, D. & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology*, 16, 353–386.

Dieckmann, N., Slovic, P., & Peters, E. (2009). The use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. *Risk Analysis*, 29, 1473–1488.

Diekman, C. & Malcolm, K. (2009). Consumer perception and insights on fats and fatty acids: Knowledge on the quality of diet fat. *Annals of Nutrition and Metabolism*, 54, 25–32.

Dieussaert, K., Verkerk, S., Gillard, E., & Schaeken, W. (2011). Some effort for some: Further evidence that scalar implicatures are effortful. *Quarterly Journal of Experimental Psychology (Hove)*, 64, 2352–67.

Donovan, R. & Jalleh, G. (1999). Positively versus negatively framed product attributes: The influence of involvement. *Psychology and Marketing*, 16, 613–630.

Druzdzel, M. J. (1989). Verbal uncertainty expressions: Literature review.

Du, N. & Stevens, K. (2011). Numeric-to-verbal translation of probability expressions in sfas 5. *Managerial Auditing Journal*, 26, 248–262.

Dunn, T., Lutes, D., & Risko, E. (2016). Metacognitive evaluation in the avoidance of demand. *Journal of Experimental Psychology: Human Perception and Performance*, 42, 1372–1387.

Dunwoody, P., Haarbauer, E., Mahan, R., Marino, C., & Tang, C.-C. (2000). Cognitive adaptation and its consequences: A test of cognitive continuum theory. *Journal of Behavioral Decision Making*, 13, 35–54.

Dwivedi, V., Goertz, K., & Selvanayagam, J. (2018). Heuristics in language comprehension. *Journal of Behavioral and Brain Science*, 8, 430–446.

Ebneter, D., Latner, J., & Nigg, C. (2013). Is less always more? the effects of low-fat labeling and caloric information on food intake, calorie estimates, taste preference, and health attributions. *Appetite*, 68, 92–97.

Edwards, A., Elwyn, G., & Mulley, A. (2002). Explaining risks: Turning numerical data into meaningful pictures. *British Medical Journal*, (pp. 324).

Einhorn, H. & Hogarth, R. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Reviews of Psychology*, 32, 53–88.

Emrich, T. E., Qi, Y., Mendoza, J. E., Lou, W., Cohen, J. E., & L'Abbe M, R. (2014). Consumer perceptions of the nutrition facts table and front-of-pack nutrition rating systems. *Applied Physiology, Nutrition and Metabolism*, 39, 417–24.

Epley, N., Mak, D., & Idson, L. (2006). Bonus or rebate?: The impact of income framing on spending and saving. *Journal of Behavioral Decision Making*, 19, 213–227.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, 71, 390–405.

Erev, I. & Cohen, B. (1990). Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. *Organizational Behavior and Human Decision Processes*, 45, 1–18.

Evans, J. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, 128, 978–996.

Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255–278.

Evans, J. & Stanovich, K. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8, 223–241.

Fritz, M. & MacKinnon, D. (2007). Required sample size to detect the mediated effect. *Psychological Science*, 18, 233–239.

Gal, D. & Rucker, D. (2018). Loss aversion, intellectual inertia, and a call for a more contrarian science: A reply to Simonson & Kivetz and Higgins & Liberman. *Journal of Consumer Psychology*, 28, 533–539.

Gamliel, E. (2013). To end life or not to prolong life: The effect of message framing on attitudes toward euthanasia. *Journal of Health Psychology*, 18, 693–703.

Gamliel, E. & Kreiner, H. (2013). Is a picture worth a thousand words? the interaction of visual display and attribute representation in attenuating framing bias. *Judgment and Decision Making*, 8, 482–491.

Gamliel, E. & Peer, E. (2010). Attribute framing affects the perceived fairness of health care allocation principles. *Judgment and Decision Making*, 5, 11–20.

Gawronski, B. & Ye, Y. (2014). What drives priming effects in the affect misattribution procedure? *Personality and Social Psychology Bulletin*, 40, 3–15.

Gigerenzer, G. (2011). Personal reflections on theory and psychology. *Theory and Psychology*, 20, 733–743.

Gigerenzer, G. (2016). Towards a rational theory of heuristics. In R. Frantz & L. Marsh (Eds.), *Minds, models, and milieux: Commemorating the centennial of the birth of Herbert Simon* (pp. 34–59). New York, NY: Palgrave Macmillan.

Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–669.

Gigerenzer, G., Hertwig, R., van den Broek, E., Fasolo, B., & Katsikopoulos, K. (2005). "a 30% chance of rain tomorrow": How does the public understand probabilistic weather forecasts? *Risk Analysis*, 25, 623–629.

Gigerenzer, G. & Kolpatzik, K. (2017). How new fact boxes are explaining medical risk to millions. *British Medical Journal*, 357, j2460.

Gigerenzer, G. & Todd, P. M. (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Gilovich, T. & Griffin, D. (2010). Judgment and decision making. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology*, volume 1 (pp. 542–588). Hoboken, NJ: John Wiley & Sons.

Glöckner, A. & Betsch, T. (2008a). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making*, 3, 215–228.

Glöckner, A. & Betsch, T. (2008b). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1055–1075.

Glöckner, A. & Betsch, T. (2011). The empirical content of theories in judgment and decision making: Shortcomings and remedies. *Judgment and Decision Making*, 6, 711–721.

Glöckner, A., Hilbig, B., & Jekel, M. (2014). What is adaptive about adaptive decision making? a parallel constraint satisfaction account. *Cognition*, 133, 641–666.

Glöckner, A. & Hodges, S. (2010). Parallel constraint satisfaction in memory-based decisions. *Experimental Psychology*, 58, 180–195.

Glöckner, A. & Witteman, C. (2010a). Beyond dual-process models: A categorisation of processes underlying intuitive judgement and decision making. *Thinking and Reasoning*, 16, 1–25.

Glöckner, A. & Witteman, C. (2010b). Foundations for tracing intuition: Models, findings, categorizations. In A. Glöckner & C. Witteman (Eds.), *Foundations for tracing intuition: Challenges and methods* (pp. 1–23). Hove, UK: Psychology Press.

Gomez, P. (2013). Common biases and heuristics in nutritional quality judgments: A qualitative exploration. *International Journal of Consumer Studies*, 37, 152–158.

González-Vallejo, C., Erev, I., & Wallsten, T. (1994). Do decision quality and preference order depend on whether probabilities are verbal or numerical? *The American Journal of Psychology*, 107, 157–172.

Graham, D. J. & Jeffery, R. (2011). Location, location, location: Eye-tracking evidence that consumers preferentially view prominently positioned nutrition information. *Journal of the American Dietary Association*, 111, 1704–1711.

Gravel, K., Doucet, E., Herman, C. P., Pomerleau, S., Bourlaud, A.-S., & Provencher, V. (2012). 'healthy', 'diet', or 'hedonic'. how nutrition claims affect food-related perceptions and intake? *Appetite*, 59, 877–884.

Grice, H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics, vol 3: Speech Arts*, volume 3 (pp. 41–58). New York, NY: Academic Press.

Grodner, D. & Sedivy, J. (2011). The effect of speaker-specific information on pragmatic inferences. In E. A. Gibson & N. J. Pearlmutter (Eds.), *The processing and acquisition of reference* (pp. 239–271). Cambridge, MA: MIT Press.

Grunert, K., Fernàndez-Celemìn, L., Wills, J., Storcksdieck genannt Bonsmann, S., & Nureeva, L. (2010a). Use and understanding of nutrition information on food labels in six european countries. *Zeitschrift Fur Gesundheitswissenschaften*, 18, 261–277.

Grunert, K. & Wills, J. (2007). A review of european research on consumer response to nutrition information on food labels. *Journal of Public Health*, 15, 385–399.

Grunert, K., Wills, J., & Fernandez-Celemin, L. (2010b). Nutrition knowledge, and use and understanding of nutrition information on food labels among consumers in the uk. *Appetite*, 55, 177–89.

Guiné, R. P. F., Duarte, J., Ferreira, M., Correia, P., Leal, M., Rumbak, I., ... Straumite, E. (2016). Attitudes towards dietary fibre on a multicultural basis: A fibre study framework. *Current Nutrition and Food Science*, 12, 132–141.

Hammond, K. (1980). The integration of research in judgement and decision theory. https://apps.dtic.mil/docs/citations/ADA088471.

Hammond, K. (1981). Principles of organization in intuitive and analytical cognition. https://apps.dtic.mil/docs/citations/ADA096570.

Hammond, K. (1988). Judgment and decision making in dynamic tasks. https://apps.dtic.mil/docs/citations/ADA199907.

Hamrosi, K., Dickinson, R., Knapp, P., Raynor, D., Krass, I., Sowter, J., & Aslani, P. (2012). It's for your benefit: exploring patients' opinions about the inclusion of textual and numerical benefit information in medicine leaflets. *International Journal of Pharmacy Practice*, 21, 216–225.

Hardisty, D., Johnson, E., & Weber, E. (2010). A dirty word or a dirty world? attribute framing, political affiliation, and query theory. *Psychological Science*, 21, 86–92.

Harris, A., Corner, A., & Hahn, U. (2009). Estimating the probability of negative events. *Cognition*, 110, 51–64.

Hawley, K. L., Roberto, C. A., Bragg, M. A., Liu, P. J., Schwartz, M. B., & Brownell, K. D. (2013). The science on front-of-package food labels. *Public Health Nutrition*, 16, 430–9.

Hayes, A. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Methodology in the Social Sciences. New York, NY: The Guilford Press.

Healy, A. (1976). Detection errors on the word the: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 235–242.

Hearty, Á. P., McCarthy, S., Kearney, J., & Gibney, M. (2007). Relationship between attitudes towards healthy eating and dietary behaviour, lifestyle and demographic factors in a representative sample of irish adults. *Appetite*, 48, 1–11.

Hersey, J. C., Wohlgenant, K. C., Arsenault, J. E., Kosa, K. M., & Muth, M. K. (2013). Effects of front-of-package and shelf nutrition labeling systems on consumers. *Nutrition Review*, 71, 1–14.

Hertwig, R. & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, (pp. 1–14).

Hieke, S. & Taylor, C. (2012). A critical review of the literature on nutritional labeling. *Journal of Consumer Affairs*, 46, 120–156.

Higginson, C., Kirk, T., Rayner, M., & Draper, S. (2002). How do consumers use nutrition label information? *Nutrition and Food Science*, 32, 145–152.

Hilton, D. (2008). Emotional tone and argumentation in risk communication. *Judgment & Decision Making*, 3, 100–110.

HM Revenue and Customs (2016). Soft drinks industry levy. `https://www.gov.uk/government/publications/soft-drinks-industry-levy/soft-drinks-industry-levy`.

Honda, H. & Yamagishi, K. (2017). Communicative functions of directional verbal probabilities: Speaker's choice, listener's inference, and reference points. *The Quarterly Journal of Experimental Psychology*, 70, 2141–2158.

Horn, L. (1989). *A natural history of negation*. Chicago, IL: University of Chicago Press.

Horn, L. (2006). Implicature. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 3–28). Maldon, MA: Blackwell Publishing Ltd.

Horn, L. & Ward, G. (2006). *The handbook of pragmatics*. Maldon, MA: Blackwell Publishing Ltd.

Hornsey, M. J., Harris, E. A., Bain, P. G., & Fielding, K. S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change*, 6, 622–626.

Horstmann, N., Ahlgrimm, A., & Glöckner, A. (2009). How distinct are intuition and deliberation? an eye-tracking analysis of instruction-induced decision modes.

Horstmann, N., Hausmann, D., & Ryf, S. (2010). Methods of inducing intuitive and deliberate processing modes. In A. Göckner & C. Witteman (Eds.), *Foundations for tracing intuition: Challenges and methods* (pp. 219–237). Hove, UK: Psychology Press.

Huizingh, E. & Vrolijk, H. (1997). A comparison of verbal and numerical judgments in the analytic hierarchy process. *Organizational Behavior and Human Decision Processes*, 70, 237–247.

Ingram, J., Hand, C., & Moxey, L. (2014). Processing inferences drawn from the logically equivalent frames half full and half empty. *Journal of Cognitive Psychology*, 26, 799–817.

Ioannidis, J. (2005). Why most published research findings are false. *PLoS Medicine*, 2, e124.

Isaac, M. & Poor, M. (2016). The sleeper framing effect: The influence of frame valence on immediate and retrospective judgments. *Journal of Consumer Psychology*, 26, 53–65.

Jaffe-Katz, A., Budescu, D., & Wallsten, T. (1989). Timed magnitude comparisons of numerical and nonnumerical expressions of uncertainty. *Memory and Cognition*, 7, 249–264.

Janiszewski, C., Silk, T., & Cooke, A. (2003). Different scales for different frames: The role of subjective scales and experience in explaining attribute-framing effects. *Journal of Consumer Research*, 30, 311–25.

Janssens, S., Chokoshvilli, D., Binst, C., Mahieu, I., Henneman, L., Paepe, A. D., & Borry, P. (2016). Attitudes of cystic fibrosis patients and parents toward carrier screening and related reproductive issues. *European Journal of Human Genetics*, 24, 506–512.

Jin, J., Zhang, W., & Chen, M. (2017). How consumers are affected by product descriptions in online shopping: Event-related potentials evidence of the attribute framing effect. *Neuroscience Research*, 125, 21–28.

John, L., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23, 524–532.

Johnson-Laird, P. N., Myers, T., Brown, K., & McGonigle, B. (1986). Reasoning without logic. In *Reasoning and discourse processes* (pp. 13–50). London, UK: Academic Press.

Juanchich, M., Sirota, M., & Bonnefon, J.-F. (2019). Verbal uncertainty. In C. Cummins & N. Katsos (Eds.), *Oxford handbook of experimental semantics and pragmatics*. Oxford University Press.

Juanchich, M., Sirota, M., & Butler, C. (2012). The perceived functions of linguistic risk quantifiers and their effect on risk, negativity perception and decision making. *Organizational Behavior and Human Decision Processes*, 118, 72–81.

Juanchich, M., Sirota, M., Karelitz, T., & Villejoubert, G. (2013). Can membership-functions capture the directionality of verbal probabilities? *Thinking & Reasoning*, 19, 231–247.

Just, D. & Wansink, B. (2014). One man's tall is another man's small: How the framing of portion size influences food choice. *Health Economics*, 23, 776–791.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York, NY: Farrer, Straus and Giroux.

Kahneman, D. & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press.

Kahneman, D. & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64, 515–526.

Keenan, E. & Paperno, D. (2012). *Handbook of quantifiers in natural language*, volume 90 of *Studies in linguistics and philosophy.* New York, NY: Springer Science+Business Media.

Keren, G. (2007). Framing, intentions, and trust-choice incompatibility. *Organizational Behavior and Human Decision Processes*, 103, 238–255.

Kess, J. (1992). *Psycholinguistics: Psychology, linguistics, and the study of natural language*, volume 86 of *Current Issues in Linguistic Theory.* Amsterdam, Netherlands: John Benjamins Publishing.

Keysar, B., Hayakawa, S., & An, S. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science.*

Kim, J., Kim, J., & Marshall, R. (2014). Search for the underlying mechanism of framing effects in multi-alternative multi-attribute decision situations. *Journal of Business Research*, 67, 378–385.

Kiss, K. & Pafel, J. (2017). Quantifier scope ambiguities. In M. Everaert & H. Riemsdijk (Eds.), *The Wiley Blackwell Companion to Syntax, Second Edition.* Hoboken, NJ: John Wiley & Sons.

Knapp, P., Gardner, P., Carrigan, N., Raynor, D., & Woolf, E. (2009a). Perceived risk of medicine side effects in users of a patient information website: A study of the use of verbal descriptors, percentages and natural frequencies. *British Journal of Health Psychology*, 14, 579–594.

Knapp, P., Gardner, P., Raynor, D., Woolf, E., & McMillan, B. (2010). Perceived risk of tamoxifen side effects: A study of the use of absolute frequencies or frequency bands, with or without verbal descriptors. *Patient Education and Counseling*, 79, 267–271.

Knapp, P., Gardner, P., & Woolf, E. (2015). Combined verbal and numerical expressions increase perceived risk of medicine side-effects: A randomized controlled trial of ema recommendations. *Health Expectations*, 19, 264–274.

Knapp, P., Raynor, D., & Berry, D. (2004). Comparison of two methods of presenting risk information to patients about the side effects of medicines. *BMJ Quality and Safety*, 13, 176–180.

Knapp, P., Raynor, D., Gardner, P., Carrigan, N., & McMillan, B. (2009b). Communicating the risk of side effects to patients: an evaluation of uk regulatory recommendations. *Drug Safety*, 32, 837–849.

Knox, J. (1969). The use of numerical probability factors in public forecasts for the prediction of precipitation occurrence. *Atmosphere*, 7, 81–96.

Kreiner, H. & Gamliel, E. (2017). The role of attention in attribute framing. *Journal of Behavioral Decision Making*.

Krishnamurthy, P., Carter, P., & Blair, E. (2001). Attribute framing and goal framing effects in health decisions. *Organisational Behavior and Human Decision Processes*, 85, 382–399.

Kruglanski, A. & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, 118, 97–109.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.

Kuvaas, B. & Selart, M. (2004). Effects of attribute framing on cognitive processing and evaluation. *Organizational Behavior & Human Decision Processes*, 95, 198–207.

Lan, X. (2003). Consumers' judgments of numerical and price information. *Journal of Product and Brand Management*, 12, 275–292.

Laws, K. (2016). Psychology, replication & beyond. *BMC Psychology*, 4, 30.

Lechner, L., Brug, J., De Vries, H., van Assema, P., & Mudde, A. (1998). Stages of change for fruit, vegetable and fat intake: consequences of misconception. *Health Education Research*, 13, 1–11.

Leong, L., McKenzie, C., Sher, S., & MÃŒller-Trede, J. (2017). The role of inference in attribute framing effects. *Journal of Behavioral Decision Making*, 10, 21–32.

Levin, I. (1987). Associative effects of information framing. *Bulletin of the Psychonomic Society*, 25, 85–86.

Levin, I. & Gaeth, G. (1988). How consumers are affected by the framing of attribute information before and after consuming the product. *Journal of Consumer Research*, 15, 374–378.

Levin, I., Schneider, S., & Gaeth, G. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, 76, 149–188.

Levy, L., Patterson, R., Kristal, A., & Li, S. (2000). How well do consumers understand percentage daily value on food labels? *American Journal of Health Promotion*, 14, 157–160.

Limb, M. (2012). Food labelling scheme agreed with retailers to remain voluntary. *British Medical Journal*, 345, e7212.

Lipkus, I., Samsa, G., & Rimer, B. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21, 37–44.

Liu, D. & Juanchich, M. (2018). Conceptual understanding and quantity inferences: A new framework for examining consumer understanding of food energy. *Public Health Nutrition*.

MacGregor, D., Slovic, P., Dreman, D., & Berry, M. (2000). Imagery, affect, and financial judgment. *Journal of Psychology and Financial Markets*, 1, 104–110.

Maciejovsky, B. & Budescu, D. (2013). Verbal and numerical consumer recommendations: Switching between recommendation formats leads to preference inconsistencies. *Journal of Experimental Psychology: Applied*, 19, 143–157.

MacLeod, A. & Pietravalle, S. (2017). Communicating risk: Variability of interpreting qualitative terms. *EPPO Bulletin*, 47, 57–68.

MacLeod, C. (1991). Half a century of research on the stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203.

Makel, M., Plucker, J., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives in Psychological Science*, 7, 537–542.

Malam, S., Clegg, S., Kirwan, S., & McGinigal, S. (2009). Comprehension and use of uk nutrition signpost labelling schemes.

Mandel, D. (2014). Do framing effects reveal irrational choice? *Journal of Experimental Psychology: General*, 143, 1185–1198.

Mandel, D. (2015). Communicating numeric quantities in context: Implications for decision science and rationality claims. *Frontiers in Psychology*, 6, 537.

Maubach, N., Hoek, J., & Mather, D. (2014). Interpretive front-of-pack nutrition labels. comparing competing recommendations. *Appetite*, 82, 67–77.

Maule, J. & Villejoubert, G. (2007). What lies beneath: Reframing framing effects. *Thinking and Reasoning*, 13, 25–44.

Maxwell, S., Lau, M., & Howard, G. (2015). Is psychology suffering from a replication crisis? *American Psychologist*, 70, 487–498.

Mazur, D., Hickam, D., & Mazur, M. (1999). How patients? preferences for risk information influence treatment choice in a case of high risk and high therapeutic uncertainty: Asymptomatic localized prostate cancer. *Medical Decision Making*, 19, 394–398.

McKenzie, C. R. & Nelson, J. (2003). What a speaker's choice of frame reveals: reference points, frame selection, and framing effects. *Psychonomic Bulletin and Review*, 10, 596–602.

Mejean, C., Macouillard, P., Peneau, S., Hercberg, S., & Castetbon, K. (2013). Consumer acceptability and understanding of front-of-pack nutrition labels. *J Hum Nutr Diet*, 26, 494–503.

Melnikoff, D. & Bargh, J. (2018). The mythical number two. *Trends in Cognitive Science*, 22, 280–293.

Miller, L. M., Cassady, D. L., Beckett, L. A., Applegate, E. A., Wilson, M. D., Gibson, T. N., & Ellwood, K. (2015). Misunderstanding of front-of-package nutrition information on us food products. *PLoS One*, 10, e0125306.

Millisecond Software (2015). Inquisit 4. `https://www.millisecond.com`.

Millisecond Software (2016). Inquisit 5. `https://www.millisecond.com`.

Morewedge, C. & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, 14, 435–440.

Moxey, L. (2006). Effects of what is expected on the focusing properties of quantifiers: A test of the presupposition-denial account. *Journal of Memory and Language*, 55, 422–439.

Moxey, L. (2017). Processing quantified noun phrases with numbers versus verbal quantifiers. *Discourse Processes*, 55, 136–145.

Moxey, L. & Sanford, A. (1986). Quantifiers and focus. *Journal of Semantics*, 5, 189–206.

Moxey, L. & Sanford, A. (1992). Context effects and the communicative functions of quantifiers: Implications for their use in attitude research. In N. Schwarz & S. Sudman (Eds.), *Context effects in social and psychological research* (pp. 279–296). New York, NY: Springer.

Moxey, L. & Sanford, A. (1993). *Communicating quantities: A psychological perspective*. Hove: Lawrence Erlbaum Associates Ltd.

Moxey, L. & Sanford, A. (2000). Communicating quantities: A review of psycholinguistic evidence of how expressions determine perspectives. *Applied Cognitive Psychology*, 14, 237–255.

Moxey, L., Sanford, A., & Dawydiak, E. (2001). Denials as controllers of negative quantifier focus. *Journal of Memory and Language*, 44, 427–442.

Munafò, M., Nosek, B., Bishop, D., Button, K., Chambers, C., du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J., & Ioannidis, J. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, 21.

Murphy, S. & Zajonc, R. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, 64, 723–739.

NatCen Social Research (2015). National diet and nutrition survey years 1-4, 2008/09-011/12.
`https://www.gov.uk/government/collections/national-diet-and-nutrition-survey`.

Newstead, S. & Collis, J. (1987). Context and the interpretation of quantifiers of frequency. *Ergonomics*, 30, 1447–1462.

Nicolas, L., Marquilly, C., & O'Mahony, M. (2010). The 9-point hedonic scale: Are words and numbers compatible? *Food Quality and Preference*, 21, 1008–1015.

Nordgren, L., Bos, M., & Dijksterhuis, A. (2011). The best of both worlds: Integrating conscious and unconscious thought best solves complex decisions. *Journal of Experimental Social Psychology*, 47, 509–511.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 2600–2606.

Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J., ...Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348, 1422–1425.

Oakes, M. (2004). Good foods gone bad: ‘infamous’ nutrients diminish perceived vitamin and mineral content of foods. *Appetite*, 42, 273–278.

Oakes, M. (2005a). Bad company: The addition of sugar, fat, or salt reduces the perceived vitamin and mineral content of foods. *Food Quality and Preference*, 16, 111–119.

Oakes, M. (2005b). Stereotypical thinking about foods and perceived capacity to promote weight gain. *Appetite*, 44, 317–324.

Olsen, A. (2015). Citizen (dis)satisfaction: An experimental equivalence framing study. *Public Administration Review*, 75, 469–478.

Olson, M. & Budescu, D. (1997). Patterns of preference for numerical and verbal probabilities. *Journal of Behavioral Decision Making*, 19, 117–131.

Open Science Collaboration (2014). Maximizing the reproducibility of your research. In S. Lilienfeld & I. Waldman (Eds.), *Psychological science under scrutiny: Recent challenges and proposed solutions*. New York, NY: Wiley.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349, 6251.

Orquin, J. & Holmqvist, K. (2017). Threats to the validity of eye-movement research in psychology. *Behavior Research Methods*.

Orquin, J. & Holmqvist, K. (2018). A primer on eye tracking methodology for behavioral sciences. In M. Schulte-Mecklenbeck, A. KÃŒhberger, & J. Johnson (Eds.), *Handbook of process tracing methods*. Routledge.

Orquin, J. & Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144, 190–206.

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin and Review*, 11, 988–1010.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford Psychology Series. New York, NY: Oxford University Press.

Parmenter, K., Waller, J., & Wardle, J. (2000). Demographic variation in nutrition knowledge in england. *Health Education Research*, 15, 163–74.

Parsons, S. & Bynner, J. (1998). Influences on adult basic skills. factors affecting the development of literacy and numeracy from birth to 37. `https://eric.ed.gov/?id=ED419160`.

Patt, A. & Schrag, D. (2003). Using specific language to describe risk and probability. *Climatic Change*, 61, 17–30.

Payne, J. & Bettman, J. (2007). Walking with the scarecrow: The information-processing approach to decision research. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of decision making* (pp. 110–132). Oxford, UK: Blackwell Publishing Ltd.

Peng, J., Jiang, Y., Miao, D., Li, R., & Xiao, W. (2013). Framing effects in medical situations: Distinctions of attribute, goal and risky choice frames. *Journal of International Medical Research*, 41, 771–6.

Pennycook, G. (2017). A perspective on the theoretical foundation of dual process models. In W. De Neys (Ed.), *Dual Process Theory 2.0* (pp. 5–27). London, UK: Routledge.

Pennycook, G., De Neys, W., Evans, J., Stanovich, K., & Thompson, V. (2018). The mythical dual process typology. *Trends in Cognitive Sciences*, 22, 667–668.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, 21, 31–35.

Peters, E., Dieckmann, N., Västfjäll, D., Mertz, C. K., Slovic, P., & Hibbard, J. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied*, 15, 213–227.

Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs*, 26, 741–8.

Piercey, M. (2009). Motivated reasoning and verbal vs. numerical probability assessment: Evidence from an accounting context. *Organizational Behavior and Human Decision Processes*, 108, 330–341.

Piñon, A. & Gambara, H. (2005). A meta-analytic review of framming effect: risky, attribute and goal framing. *Psicothema*, 17, 325–331.

Plessner, H., Betsch, C., & Betsch, T. (2008). *Intuition in judgment and decision making*. New York, NY: Lawrence Erlbaum Associates.

Plessner, H. & Czenna, S. (2008). The benefits of intuition. In H. Plessner, C. Betsch, & T. Betsch (Eds.), *Intuition in judgment and decision making* (pp. 251–266). New York, NY: Lawrence Erlbaum Associates.

Raghunathan, R., Naylor, R., & Hoyer, W. (2006). The unhealthy = tasty intuition and its effects on taste inferences, enjoyment, and choice of food products. *Journal of Marketing*, 70, 170–184.

Rayner, M., Boaz, A., & Higginson, C. (2001). Consumer use of health-related endorsements on food labels in the united kingdom and australia. *Journal of Nutrition Education*, 33, 24–30.

Rayner, M., Scarborough, P., & Williams, C. (2004). The origin of guideline daily amounts and the food standards agency's guidance on what counts as 'a lot' and 'a little'. *Public Health Nutrition*, 7, 549–56.

Reagan, R., Mosteller, F., & Youtz, C. (1989). Quantitative meanings of verbal probability expressions. *Journal of Applied Psychology*, 74, 433–442.

Renkewitz, F., Fuchs, H., & Fiedler, S. (2011). Is there evidence of publication biases in jdm research? *Judgment and Decision Making*, 6, 870–881.

Renooij, S. & Witteman, C. (1999). Talking probabilities: Communicating probabilistic information with words and numbers. *International Journal of Approximate Reasoning*, 22, 169–194.

Reyna, V. F. & Brainerd, C. (1991). Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making*, 4, 249–262.

Reyna, V. F., Nelson, W., Han, P., & Dieckmann, N. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, 135, 943–973.

Roberto, C., Bragg, M., Livingston, K., Harris, J., Thompson, J., Seaman, M., & Brownell, K. (2012). Choosing front-of-package food labelling nutritional criteria: How smart were 'smart choices'? *Public Health Nutrition*, 15, 262–7.

Roberto, C. A. & Khandpur, N. (2014). Improving the design of nutrition labels to promote healthier food choices and reasonable portion sizes. *International Journal of Obesity*, 38, S25–33.

Robitzsch, A., Grund, S., & Henke, T. (2019). miceadds r package. `https://CRAN.R-project.org/package=miceadds`.

Roe, B., Levy, A., & Derby, B. (1999). The impact of health claims on consumer search and product evaluation outcomes: Results from fda experimental data. *Journal of Public Policy and Marketing*, 18, 89–105.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638–641.

Rusou, Z., Zakay, D., & Usher, M. (2013). Pitting intuitive and analytical thinking against each other: The case of transitivity. *Psychonomic Bulletin and Review*, 20, 608–614.

Russo, J. (2011). Eye fixations as a process trace. In M. Schulte-Mecklenbeck, A. Kühberger, J. Johnson, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research: A critical review and user's guide* (pp. 43–64). Psychology Press.

Saad, G. & Gill, T. (2014). The framing effect when evaluating prospective mates: An adaptationist perspective. *Evolution and Human Behavior*, 35, 184–192.

Saikh, T., Bangalore, S., Carl, M., & Bandyopadhyay, S. (2015). Predicting source gaze fixation duration: A machine learning approach. In *2015 International Conference on Cognitive Computing and Information Processing (CCIP)*: IEEE.

Sanford, A., Dawydiak, E., & Moxey, L. (2007). A unified account of quantifer perspective effects in discourse. *Discourse Processes*, 44, 1–32.

Sanford, A., Fay, N., Stewart, A., & Moxey, L. (2002). Perspective in statements of quantity, with implications for consumer psychology. *Psychological Science*, 13, 130–134.

Sanford, A. & Moxey, L. (2003). New perspectives on the expression of quantity. *Current Directions in Psychological Science*, 12, 240–243.

Sanford, A., Moxey, L., & Paterson, K. (1994). Psychological studies of quantifiers. *Journal of Semantics*, 10, 153–170.

Sanford, A., Moxey, L., & Paterson, K. (1996). Attentional focusing with quantifiers in production and comprehension. *Memory & Cognition*, 24, 144–155.

Savoie, N., Barlow Gale, K., Harvey, K. L., Binnie, M. A., & Pasut, L. (2013). Consumer perceptions of front-of-package labelling systems and healthiness of foods. *Canadian Journal of Public Health*, 104, e359–63.

Scammon, D. (1977). ”information load” and consumers. *Journal of Consumer Research*, 4, 148–155.

Scheibehenne, B., Miesler, L., & Todd, P. M. (2007). Fast and frugal food choices: Uncovering individual decision heuristics. *Appetite*, 49, 578–589.

Scherbaum, C. & Ferreter, J. (2008). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, 12, 347–367.

Schmeltzer, C. & Hilton, D. (2014). To do or not to do? a cognitive consistency model for drawing conclusions from conditional instructions and advice. *Thinking and Reasoning*, 20, 16–50.

Schroyens, W., Schaeken, W., & Handley, S. (2003). In search of counter-examples: Deductive rationality in human reasoning. *The Quarterly Journal of Experimental Psychology*, 56A, 1129–1145.

Schulte-Mecklenbeck, M., Johnson, J., Böckenholt, U., Goldstein, D. G., Russo, J., Sullivan, N., & Willemsen, M. (2017a). Process-tracing methods in decision making: On growing up in the 70s. *Current Directions in Psychological Science*, 26, 442–450.

Schulte-Mecklenbeck, M., Kühberger, A., Gagl, B., & Hutzler, F. (2017b). Inducing thought processes: Bringing process measures and cognitive processes closer together. *Journal of Behavioral Decision Making*, 30, 1001–1013.

Schulte-Mecklenbeck, M., Sohn, M., de Bellis, E., Martin, N., & Hertwig, R. (2013). A lack of appetite for information and computation. Simple heuristics in food choice. *Appetite*, 71, 242–251.

Scott, V. & Worsley, A. (1994). Ticks, claims, tables and food groups: a comparison for nutrition labelling. *Health Promotion International*, 9, 27–37.

Sen, S. (1998). Knowledge, information mode, and the attraction effect. *Journal of Consumer Research*, 25, 64–77.

Seta, J., McCormick, M., Gallagher, P., McElroy, T., & Seta, C. (2010). Voice frequency impacts hemispheric processing of attribute frames. *Journal of Experimental Social Psychology*, 46, 1089–1092.

Shaw, M. (1982). Attending to multiple sources of information: The integration of information in decision making. *Cognitive Psychology*, 14, 353–409.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T., Cohen, J., & Botvinick, M. (2017). Towards a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, 40, 99–124.

Sher, S. & McKenzie, C. (2006). Information leakage from logically equivalent frames. *Cognition*, 101, 467–94.

Sher, S. & McKenzie, C. (2008). Framing effects and rationality. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 79–96). Oxford, UK: Oxford University Press.

Shikhare, S., Heim, S., Klein, E., Huber, S., & Willmes, K. (2015). Processing of numerical and proportional quantifiers. *Cognitive Science*, 39, 1504–1536.

Shiv, B. & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, 26, 278–292.

Siegrist, M., Leins-Hess, R., & Keller, C. (2015). Which front-of-pack nutrition label is the most efficient one? the results of an eye-tracker study. *Food Quality and Preference*, 39, 183–190.

Simon, H. (1990). Invariants of human behaviour. *Annual Review of Psychology*, 41, 1–19.

Sinclair, S., Hammond, D., & Goodman, S. (2013). Sociodemographic differences in the comprehension of nutritional labels on food products. *Journal of Nutrition Education and Behavior*, 45, 767–72.

Sink, S. (1995). Determining the public's understanding of precipitation forecasts: Results of a survey. *National Weather Digest*, 19, 9–15.

Sirota, M. & Juanchich, M. (2012). To what extent do politeness expectations shape risk perception? even numerical probabilities are under their spell! *Acta Psychologica*, 141, 391–399.

Sirota, M. & Juanchich, M. (2015). A direct and comprehensive test of two postulates of politeness theory applied to uncertainty communication. *Judgment and Decision Making*, 10, 232–240.

Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. (2007). The affect heuristic. *European Journal of Operational Research*, 177, 1333–1352.

SR Research (2007). Eyelink user manual. `http://sr-research.jp/support/EyeLink%201000%20User%20Manual%201.5.0.pdf`.

Steiger, A. & Kühberger, A. (2018). A meta-analytic re-appraisal of the framing effect. *Zeitschrift für Psychologie*, 226, 45–55.

Steptoe, A., Pollard, T. M., & Wardle, J. (1995). Development of a measure of the motives underlying the selection of food: The food choice questionnaire. *Appetite*, 25, 267–84.

Stevens, J. (2017). Replicability and reproducibility in comparative psychology. *Frontiers in Psychology*, 8, 862.

Stone, D. & Schkade, D. (1991). Numeric and linguistic information representation in multiattribute choice. *Organizational Behavior and Human Decision Processes*, 49, 42–59.

Storcksdieck genannt Bonsmann, S., Férnández Celemín, L., Larrañaga, A., Egger, S., Wills, J. M., Hodgkins, C., & Raats, M. (2010). Penetration of nutrition information on food labels across the EU-27 plus Turkey. *European Journal of Clinical Nutrition*, 64, 1379–85.

Storcksdieck genannt Bonsmann, S. & Wills, J. M. (2012). Nutrition labeling to prevent obesity: Reviewing the evidence from Europe. *Curent Obesity Research*, 1, 134–140.

Sullivan, N., Hutcherson, C., Harris, A., & Rangel, A. (2015). Dietary self-control is related to the speed with which health and taste attributes are processed. *Psychological Science*, 26, 122–134.

Szymanik, J. & Zajenkowski, M. (2009). Comprehension of simple quantifiers: Empirical evaluation of a computational model. *Cognitive Science*, 34, 521–532.

Teigen, K. & Brun, W. (1995). Yes, but it is uncertain: Direction and communicative intention of verbal probabilistic terms. *Acta Psychologica*, 88, 233–258.

Teigen, K. & Brun, W. (1999). The directionality of verbal probability expressions: Effects on decisions, predictions, and probabilistic reasoning. *Organizational Behavior and Human Decision Processes*, 80, 155–190.

Teigen, K. & Brun, W. (2000). Ambiguous probabilities: When does p = 0.3 reflect a possibility, and when does it express a doubt? *Journal of Behavioral Decision Making*, 13, 345–362.

Teigen, K. & Brun, W. (2003). Verbal probabilities: A question of frame? *Journal of Behavioral Decision Making*, 16, 53–72.

Teigen, K., Juanchich, M., & Filkukova, P. (2014). Verbal probabilities: An alternative approach. *Quarterly Journal of Experimental Psychology*, 67, 124–46.

Teigen, K. & Nikolaisen, M. (2009). Incorrect estimates and false reports: How framing modifies truth. *Thinking and Reasoning*, 15, 268–293.

Thaler, R. & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* New Haven, CT: Yale University Press.

Theil, M. (2002). The role of translations of verbal into numerical probability expressions in risk management: A meta-analysis. *Journal of Risk Research*, 5, 177–186.

Thomas, A. & Millar, P. (2012). Reducing the framing effect in older and younger adults by encouraging analytic processing. *J Gerontol B Psychol Sci Soc Sci*, 67, 139–49.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59.

Trémolière, B., De Neys, W., & Bonnefon, J.-F. (2014). The grim reasoner: Analytical reasoning under mortality salience. *Thinking and Reasoning*, 20, 333–351.

Trippas, D., Thompson, V., & Handley, S. (2017). When fast logic meets slow belief: Evidence for a parallel-processing model of belief bias. *Memory and Cognition*, 45, 539–552.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–8.

Tversky, A. & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business Ethics*, 59, S251–278.

Tversky, A. & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106, 1039–1061.

Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 166–179.

UK Department of Health (2011). Nutrition and health claims. guidance to compliance with regulation (ec) 1924/2006 on nutrition and health claims made on foods. `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/204320/Nutrition_and_health_claims_guidance_November_2011.pdf`.

UK Department of Health (2015). Food and diet. `https://www.nhs.uk/common-health-questions/food-and-diet/`.

UK Department of Health (2016). Technical guidance on nutrition labelling. https://www.gov.uk/government/publications/technical-guidance-on-nutrition-labelling.

UK Food and Drink Federation (2009). Science behind guideline daily amounts. http://www.fdf.org.uk/publications_health_wellbeing.aspx.

UK Food Standards Agency (2007). Food. using traffic lights to make healthier choices. https://www.resourcesorg.co.uk/assets/pdfs/foodtrafficlight1107.pdf.

UK Food Standards Agency (2008). Food labels. more informed choices. http://www.welllondon.org.uk/files/845/healthy-eating/14.%20Food%20Labels%20-%20More%20informed%20choices.pdf.

US Food and Drug Administration (2004). Calories count: Report of the working group on obesity. https://www.fda.gov/Food/FoodScienceResearch/ConsumerBehaviorResearch/ucm081770.htm.

van Buiten, M. & Keren, G. (2009). Speakers' choice of frame in binary choice: Effects of recommendation mode and option attractiveness. *Judgment and Decision Making*, 4, 51–63.

van Herpen, E., Hieke, S., & van Trijp, H. (2014). Inferring product healthfulness from nutrition labelling. the influence of reference points. *Appetite*, 72, 138–49.

van Herpen, E. & van Trijp, H. (2011). Front-of-pack nutrition labels. their effect on attention and choices when consumers have varying goals and time constraints. *Appetite*, 57, 148–60.

Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, 36, 1–48.

Visschers, V. (2008). Probability information in risk communication: A review of the research literature. *Risk Analysis*, 29, 267–87.

Viswanathan, M. & Childers, T. (1996). Processing of numerical and verbal product information. *Journal of Consumer Psychology*, 5, 359–385.

Viswanathan, M. & Childers, T. (1997). '5' calories or 'low' calories? what do we know about using numbers or words to describe products and where do we go from here? *Advances in Consumer Research*, 24, 412–418.

Viswanathan, M. & Narayanan, S. (1994). Comparative judgments of numerical and verbal attribute labels. *Journal of Consumer Psychology*, 3, 79–101.

Voutilainen, A., Pitkäaho, T., Kvist, T., & Vehviläinen-Julkunen, K. (2016). How to ask about patient satisfaction? the visual analogue scale is less vulnerable to confounding factors and ceiling effect than a symmetric likert scale. *Journal of Advanced Nursing*, 72, 946–957.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H., & Kievit, R. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, 7, 632–638.

Wallsten, T., Budescu, D., Zwick, R., & Kemp, S. M. (1993). Preferences and reasons for communicating probabilistic information in verbal or numerical terms. *Bulletin of the Psychonomic Society*, 31, 135–138.

Wansink, B. & Chandon, P. (2006). Can "low-fat" nutrition labels lead to obesity? *Journal of Marketing Research*, 43, 605–617.

Wansink, B. & Sobal, J. (2007). Mindless eating: The 200 daily food decisions we overlook. *Environment and Behavior*, 39, 106–123.

Watson, W., Chapman, K., King, L., Kelly, B., Hughes, C., Louie, J., Crawford, J., & Gill, T. (2013). How well do australian shoppers understand energy terms on food labels? *Public Health Nutrition*, 16, 409–417.

Weber, E. & Hilton, D. (1990). Contextual effects in the interpretations of probability words: Perceived base rate and severity of events. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 781–789.

Webster, R., Weinman, J., & Rubin, G. (2017). People's understanding of verbal risk descriptors in patient information leaflets: A cross-sectional national survey of 18- to 65-year-olds in England. *Drug Safety*, 40, 1–12.

Welkenhuysen, M., Evers-Kiebooms, G., & d' Ydewalle, G. (2001). The language of uncertainty in genetic risk communication: Framing and verbal versus numerical information. *Patient Education and Counseling*, 43, 179–187.

Wicherts, J., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS One*, 6, e26828.

Wilson, T. & Schooler, J. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, 60, 181–192.

Windschitl, P., Smith, A., Scherer, A., & Suls, J. (2017). Risk it? Direct and collateral impacts of peers' verbal expressions about hazard likelihoods. *Thinking and Reasoning*, 23, 259–291.

Windschitl, P. & Wells, G. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied*, 2, 343–364.

Wise, J. (2013). Consistent food labelling system is rolled out across UK. *BMJ*, 346.

World Health Organization (2016). Obesity and overweight: Fact sheet. `http://www.who.int/mediacentre/factsheets/fs311/en/`.

Xu, J.-H., Ye, X.-B., & Li, S. (2008). Communication mode preference paradox among native Chinese speakers. *The Journal of Social Psychology*, 149, 125.

Yong, E. (2012). Replication studies: Bad copy. *Nature*, 485, 298–300.

Zhao, M., Liu, T., Chen, G., & Chen, F. (2015). Are scalar implicatures automatically processed and different for each individual? A mismatch negativity (MMN) study. *Brain Research*, 1599, 137–149.

Zimmer, A. (1983). Verbal vs. numerical processing of subjective probabilities. In R. W. Scholz (Ed.), *Advances in Psychology*, volume 16 (pp. 159–182). London, UK: Elsevier.

Zipkin, D. A., Umscheid, C. A., Keating, N. L., Allen, E., Aung, K., Beyth, R., ...Feldstein, D. A. (2014). Evidence-based risk communication: A systematic review. *Annals of Internal Medicine*, 161, 270–280.

Zuckerman, A. & Chaiken, S. (1998). A heuristic-systematic processing analysis of the effectiveness of product warning labels. *Psychology and Marketing*, 15, 621–642.

Appendices

Appendix A: Appendix to Chapter 2

This appendix reports supplementary material to the results presented in Chapter 2 of this thesis.

A.1 Ranking of Nutrients in Terms of Healthiness and Unhealthiness (Experiment 1)

Participants ranked eight nutrients in terms of their importance in determining the healthiness and unhealthiness of food. We used the reversed mean rank to derive an overall ‘healthiness’ and ‘unhealthiness’ score for each nutrient so that higher scores reflect greater importance. The scores for all eight nutrients are shown in Table A.1. Based on these scores, the verbal and numerical GDA labels were assigned to a nutrient, as shown in Table A.2.

Table A.1. Healthiness and unhealthiness scores for 8 nutrients in Chapter 2, Experiment1.

	Score	
Nutrient	**Healthiness**	**Unhealthiness**
Minerals (e.g., vitamins)	4.73	0.98
Protein	4.64	1.95
Calories (energy)	4.25	3.41
Fibre	4.04	1.76
Sugar	2.94	5.57
Fat	2.70	5.26
Sodium (salt)	2.55	4.54
Saturates	2.14	4.54

A.2 Interpretation of All Verbal and Numerical Labels in Experiment 1

Table A.3 shows the median verbal range assigned to the 8 numerical GDA quantifiers, and the estimate of the GDA proportions made on the visual analogue scale.

Table A.4 shows the means and standard deviations for the numerical percentage attributed to the 5 verbal GDA quantifiers, and the estimate of the GDA proportions made on the visual analogue scale.

Figure A.1 shows the distribution of the visual analogue scale proportions for all 13 GDA labels in Experiment 1.

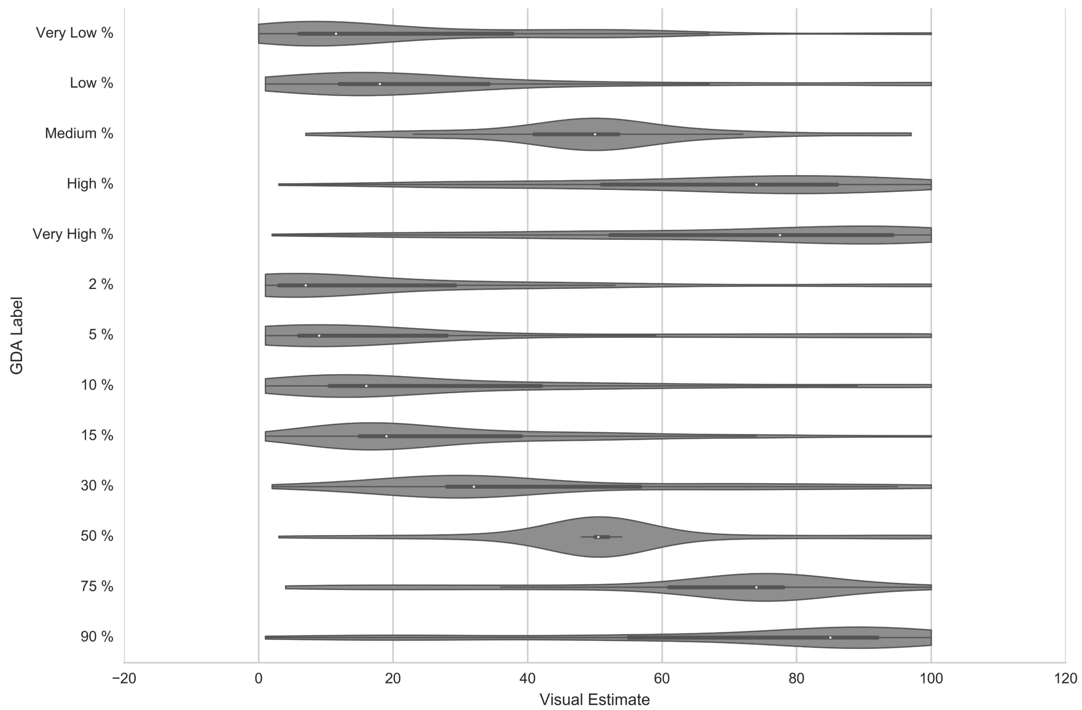

Figure A.1. Distribution of participants' visual estimates of GDA proportions for 13 labels in Chapter 2, Experiment 1.

Table A.2. Assignment of nutrient to specific quantifiers for the interpretation tasks in Chapter 2, Experiment 1).

Healthiness rank	**Interpretation task**		**Quantity perception task**	
assigned to nutrient	**Verbal**	**Numerical**	**Verbal**	**Numerical**
1	Low	5	High	30
2	High	30	Medium	15
3	Medium	15	Low	5
4	Very Low	2	Very High	50
5	Very High	50	Very Low	2
6		10		90
7		75		10
8		90		75

Table A.3. Verbal interpretation of numerical GDA quantifiers and the estimated proportion on the visual scale in Chapter 2, Experiment 1).

Numerical	**Verbal quantifier**	**Proportion (out of 100)**	
quantifier	**Median**	**Mean**	***SD***
2%	Very Low	20.96	27.34
5%	Very Low	25.90	31.20
10%	Low	29.13	27.05
15%	Low	28.84	21.55
30%	Low-Medium	41.65	24.88
50%	Medium	51.83	17.29
75%	High	65.83	23.05
90%	Very High	71.06	28.83

A.3 Results of the ANCOVA in Experiment 1

Table A.5 reports the effects of all factors on interpretations of verbal quantifiers in the ANCOVAs for Experiment 1. Table A.6 reports the effects of all factors in the interpretation of verbal quantifiers in the ANCOVA for Experiment 2.

A.4 Participant Demographics in Chapter 2, Experiment 2

Table A.7 shows the full range of socio-demographic characteristics for the sample in Experiment 2.

A.5 Pairwise Comparisons in the Repeated Measures ANOVAs in Chapter 2

Table A.8 shows the pairwise comparisons for numerical translations produced between quantities of verbal labels, and Table A.9 shows the comparisons at each quantity level between the positive and negative nutrients. Bonferroni-adjusted significance levels were used for all comparisons.

Table A.4. Numerical interpretations (%) of verbal GDA quantifiers and the estimated proportion on the visual scale in Chapter 2, Experiment 1).

Verbal quantifier	**Numerical value**		**Proportion (out of 100)**	
	Mean	***SD***	**Mean**	***SD***
Very Low %	9.36 %	12.35	23.11	24.13
Low %	16.92 %	16.10	28.20	26.85
Medium %	43.13 %	12.08	49.94	17.36
High %	68.16 %	19.62	67.65	24.56
Very High %	78.24 %	20.50	71.01	26.68

Table A.5. F- and p-values in the ANCOVAs for interpretations of verbal labels in Chapter 2, Experiment 2 ($df = 1, 59$)

Factor	**Quantity**									
	Very Low		**Low**		**Medium**		**High**		**Very High**	
	F	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Nutrient valence	0.11	.739	1.23	.271	0.01	.916	0.63	.432	1.04	.312
Frequency of label use	0.25	.621	0.53	.471	< .001	.999	2.46	.122	0.03	.855
Eating attitude	0.64	.426	0.24	.629	1.60	.211	0.28	.601	< .001	.999
BMI	3.59	.063	0.59	.445	0.52	.474	2.55	.116	0.44	.509
Gender	3.31	.074	2.64	.110	0.37	.546	0.27	.607	0.001	.977
Age	0.49	.378	1.55	.218	0.02	.887	0.37	.550	0.14	.711
Education level	2.93	.092	2.57	.114	0.34	.561	0.36	.548	0.10	.754
Ethnicity	0.25	.621	0.06	.806	0.68	.414	0.25	.617	0.24	.628
Nutrient valence × frequency of label use	0.13	.717	0.19	.666	1.20	.278	3.78	.057	0.26	.612
Nutrient valence × eating attitude	1.94	.168	0.003	.953	1.50	.226	0.48	.493	0.03	.866
Nutrient valence × BMI	2.03	.160	2.66	.108	1.17	.284	0.60	.441	1.24	.269

Table A.6. F- and p-values in the ANCOVA for interpretations of verbal labels in Chapter 2, Experiment 2

Factor	***df***	**Error *df***	***F***	***p***
Quantity	4	2,567	10.29	$< .001^{***}$
Nutrient valence	1	644	6.97	$.008^{***}$
Quantity × nutrient valence	4	2,567	2.24	.112
Frequency of label use	1	644	0.31	.581
Eating attitude	1	644	0.25	.617
BMI	1	644	0.31	.577
Gender	1	644	1.78	.183
Age	1	644	1.29	.258
Education level	1	644	0.89	.345
Ethnicity	1	644	1.66	.199
Occupation	1	644	0.25	.615
Native English-speaker	1	644	0.48	.487
Nutrient valence × frequency of label use	1	644	0.16	.691
Nutrient valence × eating attitude	1	644	0.02	.889
Nutrient valence × BMI	1	644	2.09	.149

Table A.7. Socio-demographic characteristics of the sample in Chapter 2, Experiment 2

	N	**Percentage of sample**
Age range		
18-24	54	6.8
25-34	138	17.4
35-44	147	18.5
45-54	151	19.0
55-64	148	18.6
65-74	157	19.7
Ethnicity		
White	734	92.2
Asian	34	4.3
Black	14	1.8
Mixed	10	1.3
Other	4	0.5
Employment status		
Full-time	330	41.3
Part-time	118	14.8
Self-employed	42	5.3
Student	22	2.8
Unemployed	98	12.3
Retired	190	23.8
Highest education level		
High school or equivalent	339	42.4
Degree or higher	285	35.7
Apprenticeship	35	4.4
Other Qualifications	96	12.0
No Qualifications	43	5.4
Region of residence		
East of England	71	8.9
East Midlands	58	7.3
London	93	11.7
North East	43	5.4
North West	81	10.3
South East	117	14.7
South West	69	8.7
West Midlands	67	8.4
Yorkshire and the Humber	59	7.4
Northern Ireland	14	1.8
Scotland	72	9.1
Wales	50	6.3

Table A.8. Pairwise comparisons between interpretations of verbal quantifiers in Chapter 2.

Comparison	**Mean difference (%)**	**SE**	$\boldsymbol{p}$	**95% CI for difference**
Experiment 1				
Very Low-Low	7.62	1.47	$< .001$	3.38, 11.87
Low-Medium	26.76	2.23	$< .001$	20.31, 33.21
Medium-High	24.81	2.21	$< .001$	18.41, 31.22
High-Very High	10.11	2.19	$< .001$	3.76, 16.46
Experiment 2				
Very Low-Low	3.79	0.47	$< .001$	2.48, 5.11
Low-Medium	16.47	0.57	$< .001$	14.88, 18.06
Medium-High	18.94	0.64	$< .001$	17.15, 20.72
High-Very High	8.22	0.64	$< .001$	6.42, 10.01

Table A.9. Pairwise comparisons between interpretations positive and negative nutrients quantities for each of the verbal quantifiers in Chapter 2

Quantifier	**Mean difference between positive & negative nutrients (%)**	**SE**	***p***	**95% CI for difference**	**η^2 P**
Experiment 1					
Very Low	0.39	3.02	.897	-5.62, 6.40	< .001
Low	6.68	3.88	.090	-1.06, 14.41	.04
Medium	-1.27	3.34	.705	-7.92, 5.38	.002
High	0.50	5.34	.925	-10.14, 11.15	< .001
Very High	-9.35	4.90	.061	-19.13, 0.43	.05
Experiment 2					
Very Low	5.15	1.04	< .001	3.11, 7.19	.03
Low	5.10	1.00	< .001	3.13, 7.07	.03
Medium	14.20	1.27	< .001	11.71, 16.68	.14
High	16.98	1.96	< .001	13.13, 20.83	.09
Very High	17.21	2.13	< .001	13.04, 21.39	.08

Appendix B: Appendix to Chapter 3

This appendix describes the results of the instruction manipulation for Experiment 1, and the multilevel analyses on response time for both experiments in Chapter 3. The Inquisit codes used to deliver the experiment and the experimental data are also shared on the Open Science Framework.

B.1 Instructions to be Analytical or Intuitive

After participants had completed 30 trials as reported in the main article, the second part of the experiment (30 additional trials) commenced with an instruction to make their judgements either intuitively or analytically (Schroyens et al., 2003). Participants were randomly assigned to one of the two conditions.

In the 'analysis' condition, participants were given the following instruction after a reminder about the task:

> *However, we are interested in how people reason about healthiness. Therefore, we would like you to think carefully about and analyse the reasons for making your judgements.*
> ***Please take your time to select the answer that you think is correct.***

In the 'intuitive' condition, the instruction was as follows:

> *However, we are interested in people's gut feelings about healthiness. This means that you should answer quickly based on your instincts.*
> ***Please select as fast as possible the answer that you think is correct.***

B.2 Manipulation Check

We included as a check after the second block a questionnaire to assess if participants had performed the tasks according to the given instructions. Participants reported their task performance in relation to ten adjective pairs that described intuition on one end of the 9-point scale and analysis on the other (e.g., *quickly –slowly, automatically –systematically*). A higher score on this scale reflected a greater use of analysis.

The manipulation check revealed that participants in the analytical condition reflected greater use of analysis (M = 4.97, SD = 1.18) than participants in the intuitive condition (M = 4.62, SD = 1.04). However, this difference was not significant, $t(91) = 1.55$, $p = .125$. Correlations between the manipulation check and the dependent variables in the instructed condition were also not significant.

B.3 Differences between Analytical and Intuitive Instructions

We carried out a MANOVA for the effect of task instruction and format on response times, subjective effort, and performance as preregistered. A difference score between instructed and uninstructed conditions was computed as the absolute difference between the two, with larger scores reflecting a greater discrepancy between conditions. As shown in Figure B.1, response times for the analytical instruction condition were less discrepant for the verbal than the numerical formats, and effort ratings showed a closer match between verbal intuition and numerical analysis conditions. Participants' intuitive performance with numerical formats matched their uninstructed performance better than their analytical performance with numerical formats, but their performance did not differ among instruction type for words. Across the three dependent variables, the interaction of format and instruction was not significant, $F(3, 87) = .04$, $p = .275$, $\eta^2_P = .04$. The full results of the interaction effect for each dependent variable is reported in Table B.1.

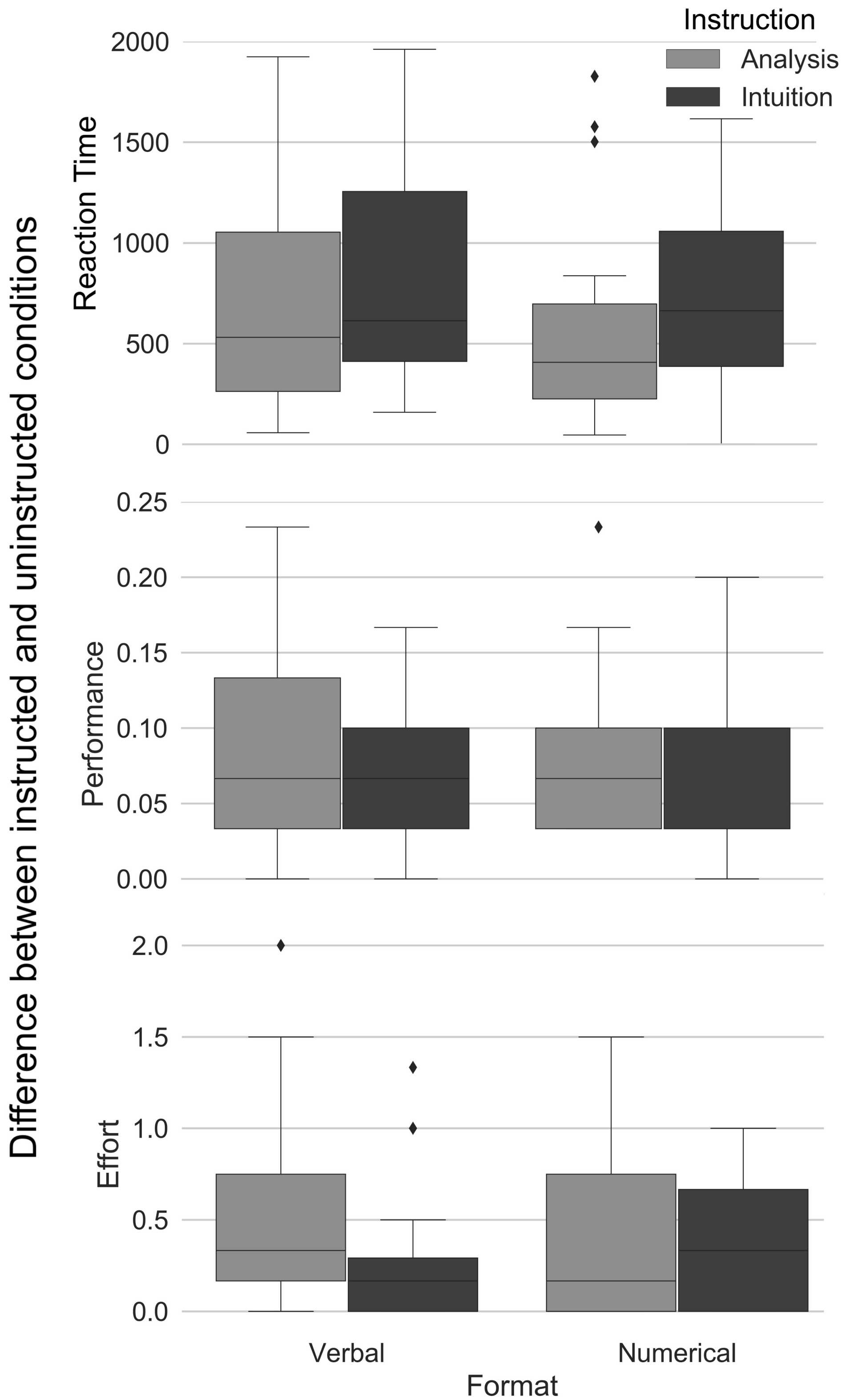

Figure B.1. Boxplots showing the distribution of the difference between uninstructed and instructed conditions in Chapter 3, Experiment 1 by format and instruction type. Central lines reflect medians and whiskers show 1.5 times the inter-quartile range.

B.4 Results of Multilevel Analysis on Response Time

Table B.2 reports the full multilevel analysis from the response time data (log-transformed) as described in Chapter 3.

Table B.1. Interaction effect of format and instruction on difference scores for response times, performance, and subjective effort in 3, Experiment 1.

	$\boldsymbol{F}$**(1, 83)**	$\boldsymbol{p}$	η^2_P
Response time	0.26	.615	.003
Performance	1.09	.299	.012
Subjective effort	1.93	.168	.021

Table B.2. Results of the multilevel analysis on response time in Experiments 1 and 2 of Chapter 3.

Factor	**Experiment 1**		**Experiment 2**	
	$\boldsymbol{F}$	$\boldsymbol{p}$	$\boldsymbol{F}$	$\boldsymbol{p}$
Format	0.55	.459	1.77	.184
Nutrient	0.49	.614	0.63	.426
Quantity	4.79	.001	0.69	.406
GDA fit	0.06	.812	0.20	.655
Format × nutrient	0.24	.787	0.03	.868
Format × quantity	2.90	.021	3.97	.047
Format × GDA fit	7.37	.007	0.44	.508
Nutrient × GDA fit	1.35	.259	2.43	.119
Quantity × GDA fit	6.83	< .001	.40	.528
Nutrient × quantity	1.11	.357	0.10	.754
Format × nutrient × quantity	0.82	.582	0.08	.782
Format × nutrient × GDA fit	0.98	.376	< .001	.985
Format × quantity × GDA fit	3.46	.008	2.94	.087
GDA Fit × nutrient × quantity	1.20	.296	1.50	.221

Appendix C: Appendix to Chapter 4

This appendix presents the results of pairwise comparisons and supplementary analyses mentioned in Chapter 4.

C.1 Pairwise Comparisons

Tables C.1 to C.3 show the pairwise comparisons for the main effects and interactions discussed in Chapter 4.

C.2 Results of a Multilevel Analysis on Response Time and Decision Performance in Experiment 1, including all Two- and Three-way Interactions

Table C.4 shows the F and p-values for all the effects in a multilevel model with all two- and three-way interactions (beyond those identified in the hypotheses) for response time and decision performance in Experiment 1.

C.3 Further Equivalence Checks for Verbal-Numerical Quantifier Interpretations

Figure C.1 shows the distribution of participants' translations for fat and minerals were significantly positively correlated at each quantity level ($r_{\text{low}} = .76$, $r_{\text{med}} = .53$, $r_{\text{high}} = .61$, $p < .001$), however the translations for low and high quantities were not significantly correlated for either nutrient ($r_{\text{fat}} = .20$, $p = .104$, $r_{\text{minerals}} = .081$, $p = .518$). Therefore, we averaged translations between the nutrients at each quantity to form an individual translation tendency for each participant. As the distributions of translations showed a slight right skew, indicating a tendency to underestimate relative to the mean, we tested in a logistic

Table C.1. Pairwise comparisons for all main effects in the multilevel models for response time (df = 2,241) and performance (df = 2,256) in Experiment 1 and 2 (df = 6,281) of Chapter 4

Response time		**Experiment 1**				**Experiment 2**			
Main effect	**Comparison**	**Mean diff (log ms)**	***p***	**95% CI Lower**	**95% CI Upper**	**Mean diff (log ms)**	***p***	**95% CI Lower**	**95% CI Upper**
Format	Numerical-Verbal	0.11	.003	0.04	0.18	0.01	.533	-0.04	0.02
Load	Easy-None	0.02	.238	-0.01	0.05				
	Hard-None	0.01	.917	-0.02	0.04	0.04	$<$ .001	0.03	0.05
	Hard-Easy	-0.01	.719	-0.04	0.02				
Nutrient	Minerals-Fat	0.01	.557	-0.01	0.02	0.02	.006	0.01	0.03
Quantity	Med-Low	0.03	.062	-0.001	0.06				
	High-Low	-0.01	.612	-0.03	0.02	0.02	.034	0.001	0.03
	Med-High	0.04	.023	0.004	0.07				
Correct response	Exceeds-Within	0.06	$<$ .001	0.04	0.07	-0.03	$<$ .001	-0.04	-0.01
Performance			**Experiment 1**				**Experiment 2**		
Main effect	**Comparison**	**Mean diff (% correct)**	***p***	**95% CI Lower**	**95% CI Upper**	**Mean diff (% correct)**	***p***	**95% CI Lower**	**95% CI Upper**
Format	Numerical-Verbal	13.9	$<$ .001	7.2	20.5	14.3	$<$ .001	11.1	17.5
Load	Easy-None	-0.3	1.00	-3.5	4.1				
	Hard-None	-1.1	1.00	-5.3	3.1	-1.1	.422	-3.8	1.6
	Hard-None	-1.4	1.00	-6.0	3.3				
Nutrient	Minerals-Fat	3.3	.057	-0.1	6.6	-5.2	.001	-8.4	-2.0
Quantity	Med-Low	1.9	.597	-5.1	8.8				
	High-Low	9.8	.006	2.2	17.4	12.1	$<$ .001	8.6	15.5
	Med-High	-7.9	.011	-14.3	-1.5				
Correct response	Exceeds-Within	-21.9	$<$ .001	-26.7	-17.0	21.8	$<$ .001	18.7	24.8

Table C.2. Pairwise comparisons for the format × load interaction for response time (df = 2,241) and performance (df = 2,256) in Experiment 1 and 2 (df = 6,281) of Chapter 4.

Response time		**Experiment 1**				**Experiment 2**			
Format	**Load comparison**	**Mean diff (log ms)**	***p***	**95% CI Lower**	**95% CI Upper**	**Mean diff (log ms)**	***p***	**95% CI Lower**	**95% CI Upper**
Verbal	Easy-None	0.02	.671	-0.02	0.06				
	Hard-None	0.01	.966	-0.03	0.04	0.05	< .001	0.03	0.06
	Hard-Easy	-0.01	.966	-0.05	0.03				
Numerical	Easy-None	0.02	.619	-0.02	0.06				
	Hard-None	0.01	.865	-0.02	0.05	0.04	< .001	0.02	0.05
	Hard-Easy	-0.01	.865	-0.04	0.03				

Performance		**Experiment 1**				**Experiment 2**			
Format	**Load comparison**	**Mean diff (% correct)**	***p***	**95% CI Lower**	**95% CI Upper**	**Mean diff (% correct)**	***p***	**95% CI Lower**	**95% CI Upper**
Verbal	Easy-None	3.4	.885	-4.3	11.1				
	Hard-None	0.1	.911	-5.7	7.7	-1.5	.454	-5.5	2.5
	Hard-Easy	-2.4	.911	-9.7	4.9				
Numerical	Easy-None	-1.2	1.00	-5.5	3.1				
	Hard-None	-1.8	1.00	-6.7	3.0	-0.7	.685	-4.1	2.7
	Hard-Easy	0.7	1.00	-5.1	3.7				

Table C.3. Pairwise comparisons for the format × nutrient × correct response interaction for response time (df = 2,241) and performance (df = 2,256) in Experiment 1 and 2 (df = 6,281) of Chapter 4.

Response time		Experiment 1				Experiment 2			
Format	**Nutrient**	**Mean response time diff (log ms) (Within –exceeds)**	***p***	**95% CI Lower**	**95% CI Upper**	**Mean response time diff (log ms) (Within –exceeds)**	***p***	**95% CI Lower**	**95% CI Upper**
Verbal	Fat	-0.02	.216	-0.06	0.01	0.01	.543	-0.02	0.03
	Minerals	-0.07	$< .001$	-0.11	-0.03	-0.07	$< .001$	-0.09	-0.05
Numerical	Fat	0.04	.025	0.01	0.08	0.01	.300	-0.01	0.04
	Minerals	0.09	$< .001$	0.05	0.13	-0.06	$< .001$	-0.08	-0.03
Performance		Experiment 1				Experiment 2			
Format	**Nutrient**	**Mean perf diff (% correct) (Within –exceeds)**	***p***	**95% CI Lower**	**95% CI Upper**	**Mean perf diff (% correct) (Within –exceeds)**	***p***	**95% CI Lower**	**95% CI Upper**
Verbal	Fat	18.1	$< .001$	10.4	25.7	8.0	.009	2.0	14.0
	Minerals	44.0	$< .001$	35.4	52.7	52.4	$< .001$	47.6	57.1
Numerical	Fat	11.3	$< .001$	4.7	18.0	3.2	.243	-2.2	8.5
	Minerals	15.5	$< .001$	8.6	22.4	35.8	$< .001$	30.7	41.0

Table C.4. Effects of format, cognitive load, nutrient, quantity, and correct response on response time and performance in 4, Experiment 1 (analysed in a multilevel model including all main effects and two- and three-way interactions).

	Response time (log ms)		**Performance (% correct)**	
	F	***p***	***F***	***p***
Main effects				
Format (verbal/numerical)	8.57	.003	17.67	< .001
Load	1.59	.204	0.42	.654
Nutrient	0.36	.547	4.01	.045
Quantity	3.90	.020	5.14	.006
Correct response	34.74	< .001	123.80	< .001
Interactions				
Format × load	0.01	.987	1.99	.138
Format × nutrient	1.14	.321	0.81	.367
Format × quantity	3.58	.028	4.74	.009
Format × correct response	0.94	.332	0.43	.517
Load × nutrient	1.14	.321	0.90	.408
Load × quantity	3.24	.012	2.19	.067
Load × correct response	0.64	.530	0.50	.604
Nutrient × quantity	0.73	.483	0.02	.980
Nutrient × correct response	5.64	.018	17.75	< .001
Quantity × correct response	21.43	< .001	33.22	< .001
Format × load × nutrient	0.60	.549	0.67	.511
Format × load × quantity	0.74	.563	1.09	.362
Format × load × correct response	0.35	.707	1.19	.306
Format × nutrient × quantity	1.96	.141	0.37	.693
Format × nutrient × correct response	< .001	.988	2.71	.100
Format × quantity × correct response	18.65	< .001	7.44	.001
Load × nutrient × quantity	1.08	.366	0.39	.816
Load × nutrient × correct response	1.81	.165	0.15	.863
Load × quantity × correct response	1.56	.182	2.16	.071
Nutrient × quantity × correct response	0.29	.751	0.19	.830

Note. The error *df* was 2,209 for response time and 2,224 for performance. The model used a variance components matrix and included random by-participant intercepts and random slopes for quantity.

regression if the individual translation tendency would interact with format to skew decisions (i.e., if people tend to underestimate the verbal quantity, they should select ‘healthy’ more often, resulting in an interaction between format and translation tendency).

The logistic regression found no significant interactions between format and translation tendency for low (b = .03, p = .120), medium (b = -.02, p = .124), or high quantities (b = -.003, p = .593).

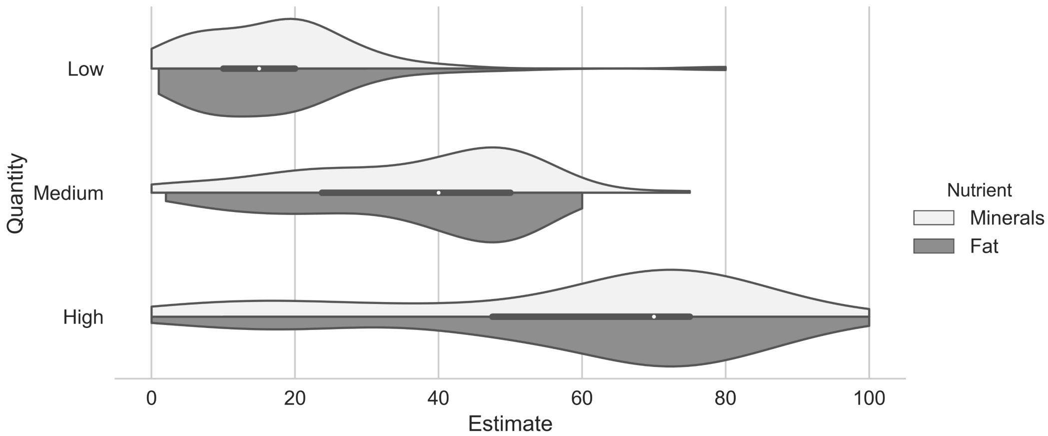

Figure C.1. Smoothed violin-plot of the distributions of participants’ translations for the verbal quantities low, medium, and high, for each nutrient.

Appendix D: Appendix to Chapter 5

This appendix presents the results of supplementary materials and analyses mentioned in Chapter 5. Figure D.1 gives examples of the four counterbalanced conditions for the food label stimuli presented to participants. Table D.1 gives the results of the multilevel analyses for number of fixations in the nutrient and quantity interest areas.

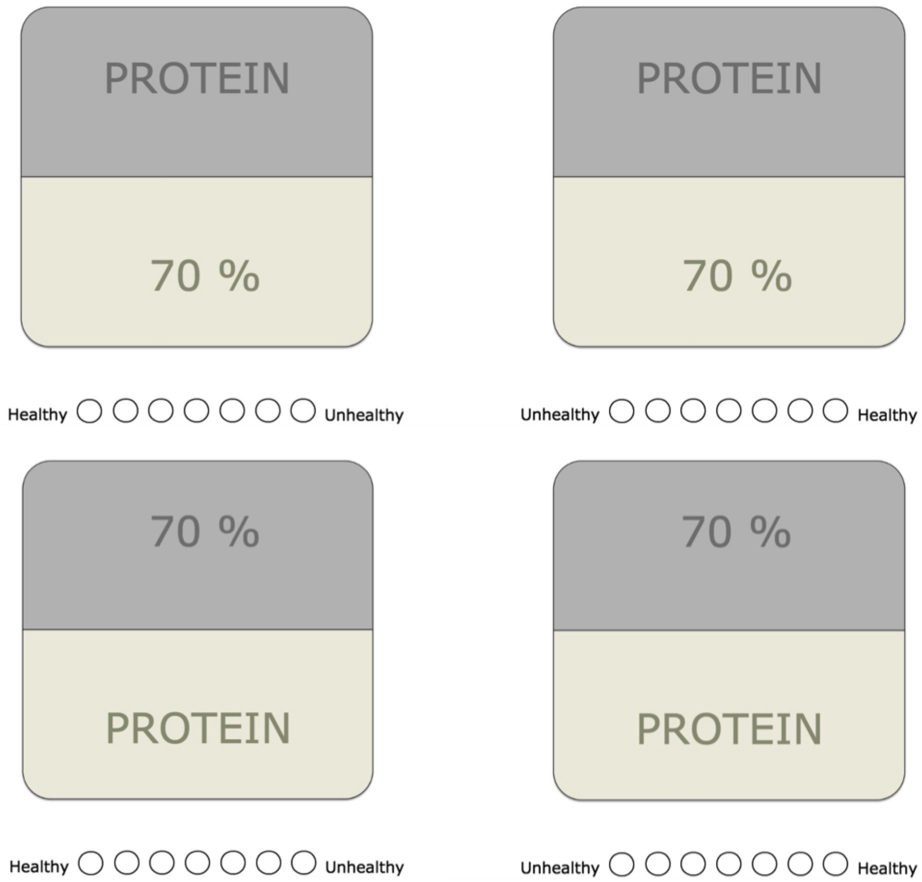

Figure D.1. Example of a numerical protein label and judgement scale shown in four counterbalanced viewing conditions. Participants were randomly assigned to one of the four viewing conditions.

Table D.1. Fixed and interaction effects for format, nutrient type, and quantity in the multilevel analyses for number of fixations on the quantifier and nutrient AOIs in Chapter 5.

Factor	***F*(error *df* = 7083)**	***df***	***p***
Number of fixations in on quantifier AOI			
Format	117.33	1	$< .001$
Nutrient valence	0.29	1	.589
Quantity	38.90	2	$< .001$
Format × nutrient valence	0.54	1	.462
Format × quantity	7.67	2	$< .001$
Nutrient valence × quantity	1.49	2	.225
Format × nutrient valence × quantity	2.29	2	.101
Number of fixations on nutrient AOI			
Format	27.46	1	$< .001$
Nutrient valence	$< .001$	1	.994
Quantity	13.01	2	$< .001$
Format × nutrient valence	.38	1	.537
Format × quantity	0.12	2	.884
Nutrient valence × quantity	2.41	2	.090
Format × nutrient valence × quantity	1.42	2	.241

Appendix E: Appendix to Chapter 6

This appendix gives the list of verbal quantifiers that participants could select from in the translation task in Chapter 6 (see Table E.1).

Table E.1. List of verbal quantifiers for translation task in Chapter 6.

Verbal quantifier	**Translation range**
Insignificant Very low Very small	Very low
Low Small	Low
Fair Medium Moderate	Medium
Large High	High
Very large Very high Significant	Very high

Appendix F: Appendix to Chapter 7

This appendix reports the supplemental analyses reported in Chapter 7, including full results of the conditional process analyses. Figure F.1 shows the effects of frame and format on healthiness judgements for the medium (40%) quantity in Experiment 1. Table F.1 shows the beta coefficients for the moderated mediation analyses of all three quantity magnitudes from Experiment 1. Table F.2 shows the beta coefficients for the moderated mediation analyses on heatlhiness and willingness to pay in Experiment 2. Table F.3 shows the beta coefficients for the moderated mediation analysis in Experiment 3. Table F.4 shows the beta coefficients for the moderated parallel mediation analysis in Experiment 4. Tables F.5 and F.6 compare the mediation and moderation effects across the four experiments.

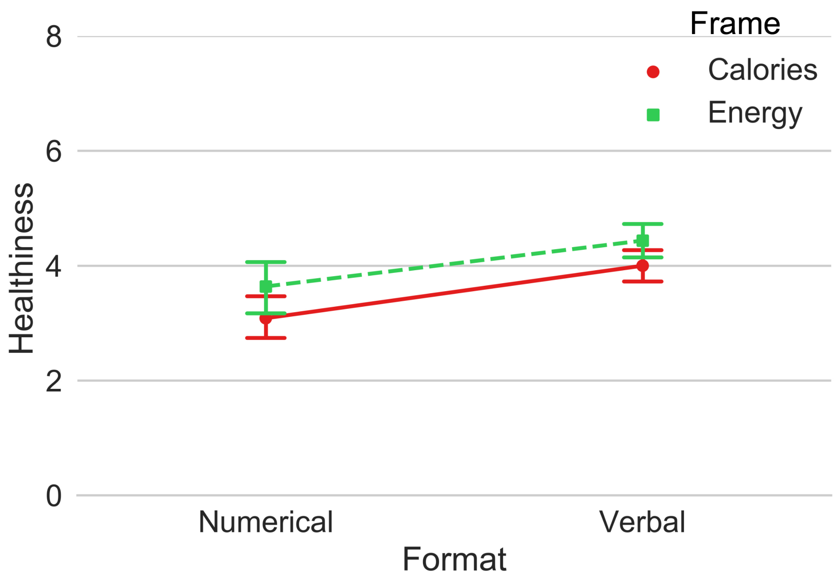

Figure F.1. Effects of frame and format on healthiness judgements of the medium (40%) quantity in Chapter 7, Experiment 1, with error bars reflecting 95% confidence intervals.

Table F.1. Effect of frame and format on affect and healthiness of 3 quantities in Chapter 7, Experiment 1 (PROCESS Model 15).

		Effect of frame on affect ($R^2 = .22$, $p < .001$)		
	b	**SE_b**	***t***	***p***
Frame (*a* path)	1.38	.19	7.30	< .001
		Effect of frame on healthiness		
Factors	***b***	**SE_b**	***t***	***p***
Quantity: Small ($R^2 = .12$, $p = .002$				
Frame (*c*' path)	-0.18	.21	-0.85	.396
Affect (*b* path)	-0.25	.08	-3.09	.002
Format	0.05	.19	0.25	.804
Affect × format	0.02	.16	0.11	.911
Frame × format	0.76	.42	1.78	.076
Quantity: Medium ($R^2 = .15$, $p < .001$)				
Frame (*c*' path)	0.35	.21	1.68	.094
Affect (*b* path)	0.11	.07	1.50	.135
Format	-0.81	.19	-4.31	< .001
Affect × format	-0.02	.15	-0.13	.898
Frame × format	0.08	.41	0.20	.846
Quantity: Large ($R^2 = .29$, $p < .001$)				
Frame (*c*' path)	0.97	.24	3.98	< .001
Affect (*b* path)	0.31	.09	3.50	.001
Format	-1.11	.24	-4.65	< .001
Affect × format	-0.24	.18	-1.36	.175
Frame × format	-0.06	.49	-0.13	.898

Note. The top panel of Figure 7.1 (see Chapter 7) illustrates the corresponding pathways for each of the effects in the model. The *a* path is the direct effect of frame on affect; the *b* path is the direct effect of affect on healthiness; the *c*' path is the direct effect of frame on healthiness after accounting for the mediated pathway.

Table F.2. Effect of frame and format on affect, healthiness, and willingness-to-pay for small and large quantities in Chapter 7, Experiment 2 (PROCESS Model 15).

		Effect of frame on affect (R^2 = .24, p < .001)		
	b	**SE$_b$**	***t***	***p***
Frame (*a* path)	1.49	.15	10.09	< .001
		Effect on healthiness		
	b	**SE$_b$**	***t***	***p***
Quantity: Small (R^2 = .10, p < .001)				
Frame (*c*' path)	-0.43	.19	-2.29	.023
Affect (*b* path)	-0.09	.07	-1.19	.235
Format	0.14	.16	0.92	.359
Affect × format	0.07	.15	0.46	.647
Frame × format	0.95	.38	2.51	.013
Quantity: Large (R^2 = .19, p < .001)				
Frame (*c*' path)	0.63	.23	2.78	.006
Affect (*b* path)	0.39	.07	5.44	< .001
Format	-0.15	.18	-0.84	.402
Affect × format	0.02	.14	0.15	.880
Frame × format	0.05	.45	0.12	.906
		Effect on willingness-to-pay		
	b	**SE$_b$**	***t***	***p***
Quantity: Small (R^2 = .02, p = .059)				
Frame (*c*' path)	-0.18	.14	-1.28	.202
Affect (*b* path)	0.05	.05	0.99	.325
Format	0.31	.13	2.49	.013
Affect × format	-0.01	.10	-0.15	.883
Frame × format	-0.12	.28	-0.44	.660
Quantity: Large (R^2 = .05, p = .006)				
Frame (*c*' path)	-0.11	.15	-0.69	.489
Affect (*b* path)	0.17	.05	3.35	< .001
Format	0.17	.12	1.35	.177
Affect × format	0.03	.10	0.25	.804
Frame × format	-0.27	.30	-0.88	.379

Note. The top panel of Figure 7.1 (see Chapter 7) illustrates the corresponding pathways for each of the effects in the model. The *a* path is the direct effect of frame on affect; the *b* path is the direct effect of affect on healthiness; the *c*' path is the direct effect of frame on healthiness after accounting for the mediated pathway.

Table F.3. Effect of frame and format on inferred recommendations and healthiness for small and large quantities in Chapter 7, Experiment 3 (PROCESS Model 8).

		Effect of frame on recommendation		
Factors	***b***	**SE_b**	***t***	***p***
Quantity: Small (R^2 = .13, p < .001)				
Frame (*a* path)	-0.82	.13	-6.18	< .001
Format	0.17	.13	-1.29	.198
Frame × format	0.48	.26	1.83	.069
Quantity: Large (R^2 = .13, p < .001)				
Frame (*a* path)	0.86	.14	-6.17	< .001
Format	0.14	.14	0.97	.335
Frame × format	-0.42	.28	-1.50	.135
		Effect on healthiness (*b* and *c*' paths)		
Factors	***b***	**SE_b**	***t***	***p***
Quantity: Small (R^2 = .43, p < .001)				
Frame (*c*' path)	-0.40	.13	-3.07	.002
Recommendation (*b* path)	0.63	.06 1	0.70	< .001
Format	-0.13	.12	-1.08	.282
Frame × format	0.20	0.24	.85	.394
Quantity: Large (R^2 = .45, p < .001)				
Frame (*c*' path)	0.56	.16	3.40	< .001
Recommendation (*b* path)	0.76	.06	11.81	< .001
Format	0.01	.14	0.05	.958
Frame × format	0.09	.29	0.32	.750

Note. The bottom panel of Figure 7.1 (see Chapter 7) illustrates the corresponding pathways for each of the effects in the model. The *a* path is the direct effect of frame on affect; the *b* path is the direct effect of affect on healthiness; the *c*' path is the direct effect of frame on healthiness after accounting for the mediated pathway.

Table F.4. Effect of frame and format on affect, inferred recommendations, and healthiness for small and large quantities in Chapter 7, Experiment 4 (PROCESS Model 8).

		Effect of frame on affect		
Factors	***b***	***SE$_b$***	***t***	***p***
Quantity: Small R^2 = .12, p < .001)				
Frame (*a* path)	-0.90	.13	-7.20	< .001
Format	-0.08	.13	-0.66	.510
Frame × format	0.07	.25	0.29	.773
Quantity: Large R^2 = .14, p < .001)				
Frame (*a* path)	1.21	.12	9.84	< .001
Format	0.07	.12	0.59	.554
Frame × format	0.24	.25	0.99	.325
		Effect of frame on recommendation		
Factors	***b***	***SE$_b$***	***t***	***p***
Quantity: Small (R^2 = .14, p < .001)				
Frame (*a* path)	-0.94	.12	-8.04	< .001
Format	0.09	.12	0.78	.439
Frame × format	0.26	.23	1.11	.267
Quantity: Large (R^2 = .17, p < .001)				
Frame	1.15	.13	9.09	< .001
Format	0.14	.13	1.16	.247
Frame × format	0.10	.25	0.38	.702
		Effect on healthiness		
Factors	***b***	***SE$_b$***	***t***	***p***
Quantity: Small (R^2 = .53, p < .001)				
Frame (*c*' path)	-0.29	.10	-2.77	.006
Affect (*b* path)	0.46	.06	8.07	< .001
Recommendation (*b* path)	0.33	.06	5.91	< .001
Format	-0.16	.09	-1.71	.087
Frame × format	-0.16	.19	-0.82	.412
Quantity: Large (R^2 = .64, p < .001)				
Frame (*c*' path)	0.28	.11	2.45	.015
Affect (*b* path)	0.64	.07	9.80	< .001
Recommendation (*b* path)	0.33	.06	5.73	< .001
Format	-0.02	.10	-0.23	.818
Frame × format	-0.20	.19	-1.01	.311

Note. The top panel of Figure 7.4 (see Chapter 7) illustrates the corresponding pathways for each of the effects in the model. The *a* path is the direct effect of frame on affect; the *b* path is the direct effect of affect on healthiness; the *c*' path is the direct effect of frame on healthiness after accounting for the mediated pathway.

Table F.5. Comparison of mediation effects across Experiments 1-4 in Chapter 7

	Conditional direct effects on healthiness (moderated *c*' path)			
	$b_{\textbf{verbal}}$	**95% CI**	$b_{\textbf{numerical}}$	**95% CI**
Experiment 1				
Small	-0.55	[-0.01, 1.11]	0.20	[-0.42, 0.82]
Medium	0.31	[-0.18, 0.79]	0.39	[-0.27, 1.04]
High	1.00	[0.34, 1.67]	0.94	[0.24, 1.64]
Experiment 2 (Healthiness)				
Small	-0.91	[-1.42, -0.41]	0.03	[-0.51, 0.58]
Large	0.61	[0.03, 1.18]	0.66	[-0.03, 1.34]
Experiment 2 (Willingness-to-pay)				
Small	-0.12	[-0.42, 0.19]	-0.24	[-0.69, 0.22]
Large	0.03	[-0.34, 0.40]	-0.24	[-0.71, 0.23]
Experiment 3				
Small	-0.50	[-0.83, -0.16]	-0.29	[-0.65, 0.06]
Large	0.51	[0.08, 0.94]	0.60	[0.18, 1.02]
Experiment 4				
Small	-0.21	[-0.46, 0.04]	-0.37	[-0.67, -0.06]
Large	0.38	[0.09, 0.66]	0.18	[-0.12, 0.49]
	Conditional indirect effect of affect (moderated *c* path)			
	$b_{\textbf{verbal}}$	**95% CI**	$b_{\textbf{numerical}}$	**95% CI**
Experiment 1				
Small	-0.35	[-0.71, -0.01]	-0.33	[-0.64, -0.07]
Medium	0.17	[-0.10, 0.49]	0.14	[-0.16, 0.42]
Large	0.60	[0.26, 1.06]	0.26	[-0.07, 0.60]
Experiment 2 (Healthiness)				
Small	-0.18	[-0.47, 0.07]	-0.08	[-0.42, 0.24]
Large	0.56	[0.28, 0.89]	0.60	[0.28, 0.97]
Experiment 2 (Willingness-to-pay)				
Small	0.08	[-0.05, 0.28]	0.06	[-0.18, 0.31]
Large	0.23	[0.06, 0.45]	0.27	[0.06, 0.52]
Experiment 4				
Small	-0.43	[-0.64, -0.27]	-0.40	[-0.62, -0.22]
Large	0.70	[0.47, 0.97]	0.86	[0.61, 1.14]
	Conditional indirect effect of recommendation (moderated *c* path)			
	$b_{\textbf{verbal}}$	**95% CI**	$b_{\textbf{numerical}}$	**95% CI**
Experiment3				
Small	-0.67	[-0.95, -0.43]	-0.36	[-0.62, -0.13]
Large	0.82	[0.51, 1.14]	0.50	[0.19, 0.84]
Experiment 4				
Small	-0.36	[-0.53, -0.23]	-0.27	[-0.44, -0.15]
Large	0.36	[0.23, 0.56]	0.40	[0.24, 0.60]

Table F.6. Comparison of moderated mediation effects across Experiments 1-4 in Chapter 7

	Moderation of indirect effect (Test for moderated mediation)	
	Index of moderated mediation	**95% CI**
Experiment 1 (Affect)		
Small	0.02	[-0.43, 0.44]
Medium	-0.03	[-0.45, 0.34]
Large	-0.33	[-0.91, 0.13]
Experiment 2 (Affect on healthiness)		
Small	0.10	[-0.31, 0.54]
Large	0.03	[-0.36, 0.46]
Experiment 2 (Affect on willingness-to-pay)		
Small	-0.02	[-0.31, 0.26]
Large	0.04	[-0.26, 0.33]
Experiment 3 (Recommendation)		
Small	0.31	[-0.02, 0.65]
Large	-0.32	[-0.76, 0.07]
Experiment 4 (Affect)		
Small	0.03	[-0.20, 0.26]
Large	0.16	[-0.15, 0.46]
Experiment 4 (Recommendation)		
Small	0.09	[-0.07, 0.25]
Large	0.03	[-0.13, 0.20]