Note. This is an accepted manuscript for an article to be published in the journal *Thinking and Reasoning*; the current version might differ slightly from the published version.

Differences Between Decisions Made Using Verbal Or Numerical Quantifiers

Dawn Liu[1], Marie Juanchich[1], Miroslav Sirota[1], & Sheina Orbell[1]

[1]*Department of Psychology*
University of Essex
Wivenhoe Park
Colchester, CO4 3SQ
United Kingdom

All correspondence regarding this article should be addressed to the first author at dliuxi@essex.ac.uk / dawn.liu@ymail.com[2]

[2]*Present address:*
University of Essex
Wivenhoe Park
Colchester, CO4 3SQ
United Kingdom

Data and materials from the studies are available on the Open Science Framework at https://osf.io/95h46/

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.

Keywords: *verbal quantifiers; numerical quantifiers; processing styles; dual-process theory; decision context*

Differences Between Decisions Made Using Verbal or Numerical Quantifiers

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 often requiring more effort to process than words (Childers & Viswanathan, 2000). 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). This idea that numerical formats are more effortful than verbal ones drives the use of verbal formats to communicate quantities in many different contexts (e.g., nutrition information: Malam, Clegg, Kirwan, & McGinigal, 2009; medical risks: Berry, Raynor, & Knapp, 2003). However, this hypothesis has not been directly tested, and 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 in a quantitative task.

A framework for organising differences between verbal and numerical quantifiers

The postulate that verbal quantifiers are quicker and easier to understand than numerical quantifiers can be understood within the framework of 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, De Neys, Evans, Stanovich, & Thompson, 2018), describes two processing styles that typically differ in terms of consciousness, automaticity, and the amount of cognitive effort involved (for an overview of dual-process theories, see Evans, 2008, and De Neys, 2017b; for alternative and more critical views of intuitive and analytical processing, see Betsch & Glöckner, 2010, and Melnikoff & Bargh, 2018). Dual-process theory suggests several measures that should distinguish processing style differences: effort, time, and accuracy. Intuitive processing is automatic, unconscious, and quick, generating affective cognitions such as a feeling of rightness about the decision—this tends to produce judgements and decisions with less effort, in less time, and with greater affective biases, which reduce accuracy. On the other hand, analytical processing is conscious, effortful, and operates more slowly and deliberately—which should produce more accurate judgements and decisions. Within this framework, processing verbal quantifiers would be more intuitive, and numerical quantifiers more analytical (Windschitl & Wells, 1996).

If verbal quantifiers do indeed produce more intuitive processing than numerical, one would expect the measures of effort, time, and accuracy to consistently point to quicker, less effortful, and more biased judgements and decisions for verbal quantifiers. People's preferences in communicating and evaluating quantities are in line with this expectation. From the communicator's perspective, people preferred to deliver information using verbal (as opposed to numerical) quantifiers (Olson & Budescu, 1997), which have been argued to be more natural in communication (Zimmer, 1983). From the recipient's perspective, verbal quantifiers aided people in making rapid evaluations of products: 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 reported that verbal expressions of food quantities were easier to use than numerical ones in decision-making (Malam et al., 2009). In contrast, people reported that numerical information about medical decisions was cognitively effortful and difficult to understand (Peters, Hibbard, Slovic, & Dieckmann, 2007).

In addition, studies of verbal probabilities show that they differ from numerical probabilities along other dimensions. Aside from 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, Juanchich, & Filkukova, 2014). On the other hand, 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). Verbal quantifiers may be more suited to integration with context (Moxey & Sanford, 1993). This could explain why it takes less time to understand where verbal quantifiers lie in an evaluative context compared to numerical ones (e.g., whether 'high protein' vs. '30% protein' is good or bad; Viswanathan & Childers, 1997).

However, in contexts that do not rely on evaluation of the quantifier, evidence for processing differences—in terms of effort, time, and accuracy—between verbal and numerical quantifiers is mixed, and varies across tasks (e.g., Childers & Viswanathan, 2000; González-Vallejo et al., 1994; Jaffe-Katz et al., 1989; Viswanathan & Narayanan, 1994). In tasks where participants had to compare the magnitudes of two numerical or two verbal quantifiers, people were even quicker with the numerical quantifiers (Jaffe-Katz, Budescu, & Wallsten, 1989; Viswanathan & Narayanan, 1994).

Other tasks have found inconclusive evidence for speed and accuracy of decisions: in decision tasks where people selected gambles in which the chance of winning was described either with verbal or numerical probabilities (e.g., ‘likely’ vs. ‘60% chance’), people’s aggregated response times and decision performance did not differ between verbal and numerical formats (González-Vallejo et al.,1994). However, people selected gambles with verbal probabilities more when the gambles paid better; in contrast, they selected gambles with numerical probabilities more when the gamble was more likely to succeed. 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.

The lack of consistent evidence across tasks shows that the structure of a task could affect how intuitive one can be with verbal or numerical quantifiers. How well people perform on tasks can often be a case of whether they apply an intuitive style to an intuitive task, or an analytical style to an analytical task (Ayal, Rusou, Zakay, & Hochman, 2015). People find verbal quantifiers easier with evaluation tasks that often require subjective judgements and rely on affective information or preference-based judgements—verbal quantifiers were also preferred for these types of tasks (Nicolas, Marquilly, & O’Mahony, 2010; Wallsten, Budescu, Zwick, & Kemp, 1993; Wilson & Schooler, 1991). Such features are also often associated with intuitive tasks (Hammond, 1988). This suggests that verbal quantifiers may be suited to tasks that tap on their ability to generate intuitive understanding about how the quantity modifies its context. In contrast, verbal quantifiers lose their advantage on tasks that do not require such understanding, such as comparing quantifier magnitudes (Viswanathan & Narayanan, 1994).

Conversely, in tasks that involve analytical judgements and objective values, people display a preference for numerical quantifiers (Budescu & Wallsten, 1990; Juanchich & Sirota, 2019). Tasks that further require the dissociation of quantifiers from their context may also prove more suitable for numerical than verbal quantifiers (Moxey & Sanford, 1993). This may be because verbal quantifiers have variable interpretations that are closely related to the context in which people encounter them. For instance, Budescu, Por, & Broomell (2012) found that even when told official verbal-numerical descriptors for how likely a climate event would occur, participants continued to interpret verbal probabilities according to their own interpretations. In a task where one should follow an ideal criterion to reach a decision, verbal quantifiers may therefore be prone to decision biases. Attempting such tasks based on intuition is often associated with using contextual knowledge to substitute a more mentally available concept to answer a difficult question (Kahneman & Frederick, 2002).

Some evidence suggests that decisions with verbal quantifiers are more susceptible to the influence of contextual information (González-Vallejo et al., 1994; Windschitl & Wells, 1996). For instance, people tended to pick gambles with higher pay outcomes for verbal than numerical quantifiers, suggesting that these outcomes influenced the decision more (González-Vallejo et al., 1994). Further work also demonstrated that verbal quantifiers produced more judgement biases than numerical ones. For example, participants perceived a ‘1 in 10’ to be less likely than a ‘10 in 100’ chance when describing these chances with verbal, but not numerical, probabilities (Windschitl & Wells, 1996). However, one should consider whether it was simply the use of numerical information in the context that created the task incompatibility (numerical-numerical vs. numerical-verbal), rather than incompatibilities in the processing approach to the task. Therefore, to test whether verbal quantifiers encourage people to rely on contextual information more, an investigation needs to use a task where only the quantifier, and not the contextual information, can take numerical format.

The present work

Our investigation aimed to test directly the effect of format on four indicators of processing styles, while overcoming methodological issues mentioned above. Using 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., substituting contextual knowledge) that may hinder decision-making performance (Kahneman, 2011). We therefore expected that participants would process 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 more on the contextual information (Experiment 2).

In addition, we extended previous research to a novel context, nutrition communication. We chose this context because it fulfils three important criteria. First, it allowed us to design a task where associated information (e.g., a nutrient) was not also expressed as a number, in contrast to previous work where the outcome was numerical: 'you can win $20' (e.g., González-Vallejo et al., 1994; Windschitl & Wells, 1996). This meant that only the quantifier of interest (the *amount* of the nutrient) took numerical (or verbal) format. Second, nutrient quantities are commonly expressed using verbal as well as numerical formats (Malam et al., 2009). As such, there was precedent for both formats being used in the real world. Finally, the nutrients provided contextual information that could 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 quantifier. In the task, participants first saw information presented in a pictorial format, which they then had to combine with either the verbal or numerical quantifier. The structure of this task allowed participants the opportunity to make their decision based on the rules of integration, or based on their positive or negative associations about the nutrient.

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; https://osf.io/95h46).

Experiment 1

Method

Participants

The study was powered to detect a small-to-medium effect in a mixed ANOVA (Cohen's $f = .18$, $\alpha = .05$) 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 (C. Betsch, 2004), eating attitudes (Steptoe, Pollard, & Wardle, 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 2.56 (*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) × 3 (nutrient) × 5 (quantity) × 2 (correct decision) mixed design.

Materials

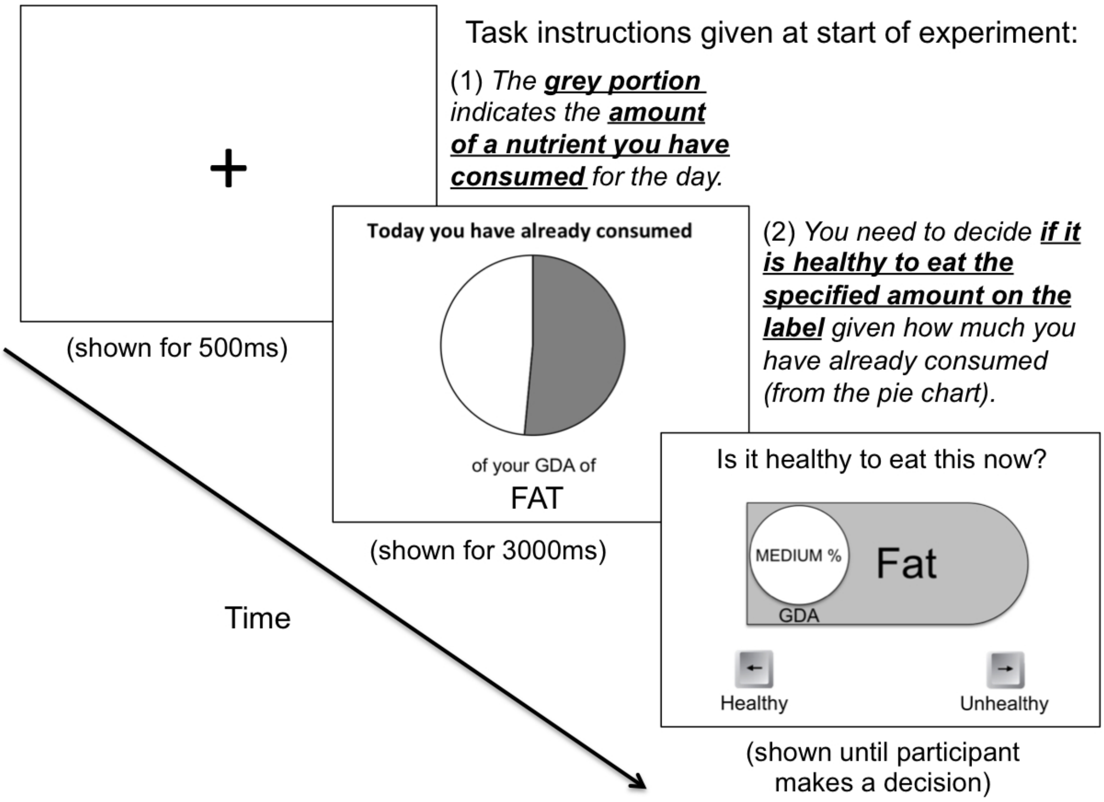
We created a decision task programmed using Inquisit 4 (Millisecond Software, 2015; code available on the OSF). In this task, participants decided whether a given standardised percentage of a nutrient was healthy—i.e., whether it could be eaten without exceeding their guideline daily amount (GDA). The 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:

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.

GDAs are the most common presentation of nutrition information in the UK and the EU (Storcksdieck genannt Bonsmann et al., 2010). Because GDAs are usually calculated based on dietary requirements for a typical person (Rayner, Scarborough, & Williams, 2004), people who think they need more than average could perceive their GDA as higher than stated, while people who think they need less than average could perceive their GDA as lower than stated. To circumvent this challenge, we also included an instruction that participants should assume the GDAs in the task were tailored to their own dietary needs. Participants then saw a pie chart and food label example, with instructions on how to complete the task (see Figure 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 full 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 1.

[Figure 1 near here. Caption: *Figure 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.]



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. 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 also gave a level of precision between a precise numerical format and a 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 combine it with vague verbal quantifiers.

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, Evers-Kiebooms, & d'Ydewalle, 2001), we matched numerical quantifiers to verbal ones (columns 3 and 4 of Table 1). To do so, we used the average translations for verbal expressions of nutrient quantities found in the first study of Liu, Juanchich, Sirota, and Orbell (2019), which used a similar sample to the present study. We did not rely on official 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 in different domains (Berry, 2006; Berry, Knapp, & Raynor, 2002; Budescu, Por, Broomell, & Smithson, 2014; Knapp, Gardner, Carrigan, Raynor, & Woolf, 2009; Knapp, Gardner, Raynor, Woolf, & McMillan, 2010; Liu et al., 2019).

Decision. Participants decided whether consuming the pie chart amount and the label amount would fall within their GDA limit (was healthy) or not (was unhealthy). The healthy button was either the left or the right arrow key. The task was in essence a mathematical one (adding two quantities and deciding whether they exceeded a third), so it could prime participants to rely more on analytical processes (Hammond, 1980). To avoid having participants all rely on analysis to solve the problem, we chose to frame the task as deciding whether a quantity was healthy or not, rather than whether the sum exceeded a quantity. This allowed a mix of both intuitive and analytical processes. Participants could give an intuitive answer based on whether the nutrient in itself was healthy or unhealthy, 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 completed at least three practice trials in which they received feedback prior to starting the experimental trials. If they incorrectly judged a within-limits quantity as unhealthy, they were informed: '*Your GDA is 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.

Response time. We measured how long participants took to make their decision from the moment they saw the label quantity. 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.

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 1), as measured in

Experiment 1 of Liu et al. (2019), which used a similar sample. The healthy (within limits) pie chart magnitude was computed as 100% - ($M_{\text{verbal quantifier translation}}$ + 1 $SD_{\text{verbal quantifier translation}}$) and the unhealthy (exceeding limits) pie chart magnitude was computed as 100% - ($M_{\text{verbal quantifier translation}}$ - 1 $SD_{\text{verbal quantifier translation}}$) (exceeds limits). 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).

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*).

[Table 1 near here.]

Table 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)
66.98 %	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 Liu et al. (2019).

Procedure

After giving informed consent and reading the instructions, participants performed a training decision block before the experimental phase. 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, Schaeken, & Handley, 2003):

Intuitive instruction:

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.

Analytical instruction:

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.

We did not pursue these analyses as the instruction manipulation was unsuccessful. We had expected that participants would spontaneously make decisions with verbal quantifiers that were akin to their decisions when told to be intuitive, and that participants would spontaneously make decisions with numerical quantifiers that were akin to their decisions 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). As such, we have 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 as a supplement to this manuscript. Data from these trials is also archived on the OSF.

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 2). On average, the numerical formats showed a trend for better performance, quicker decisions, and more effort required than verbal ones. We ran a pre-registered multivariate analysis of variance to test our hypotheses that format would have an overall effect on mean response time, task performance, and effort ratings. The analysis 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, $p > .222$).

[Table 2 near here.]

Table 2.

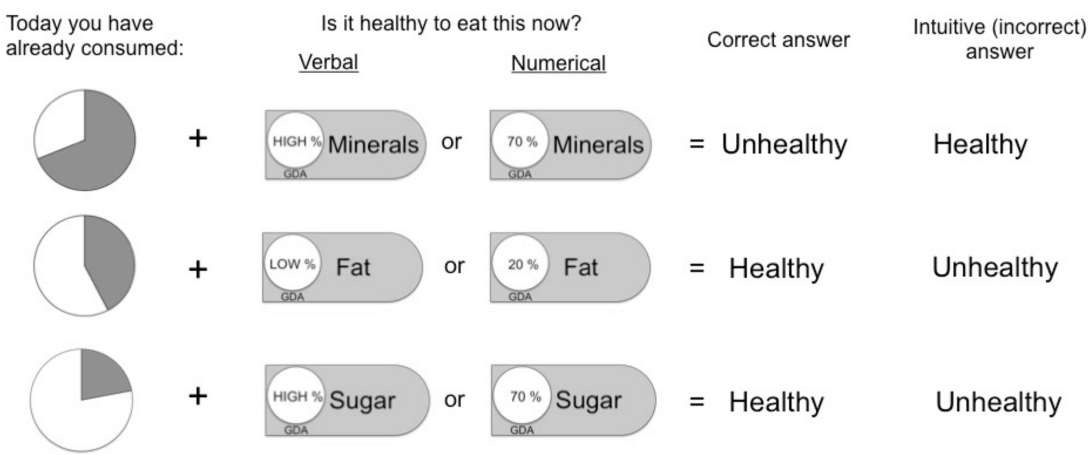
Descriptive summary of response times (ms), decision performance (% of correct trials), and effort (rating from 1: very hard to 5: very easy) across experiments and quantifier formats.

	Experiment 1		**Experiment 2**	
	Verbal	**Numerical**	**Verbal**	**Numerical**
Response time (raw, 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 2). As shown in Table 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 2 near here. Caption: *Figure 2.* Examples of nutrient and quantity combinations that had an incorrect intuitive answer.]



[Table 3 near here.]

Table 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 2).

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 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 4.

[Table 4 near here.]

Table 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 decision	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 decision	**3.20**	**.041**	**0.37**	**.544**
Format × quantity × correct decision	1.84	.119	0.49	.486
Correct decision × 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).

Participants used more context-based shortcuts with verbal quantifiers. The significant two- and three-way interactions between format, nutrient, and correct decision 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 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, $F(1, 2733) = 85.53$, $p < .001$, $M_{\text{diff}} = 39.52\%$, 95% CI [31.11, 47.90]. They did not make the same error pattern in the numerical condition, suggesting that the nutrient did not factor heavily into their decision, $F(1, 2733) = 2.54$, $p = .111$, $M_{\text{diff}} = 6.84\%$, 95% CI [-1.57, 15.25].

Discussion

Our exploratory analysis of 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 systematically interpreting verbal labels as less than the values used in the numerical condition. 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 interpersonal variation.

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 using multilevel modelling, addressing the issue of potentially inflated Type I error rates when relying on exploratory analyses (Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 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.

Method

Participants

The experiment was powered to detect the format × nutrient × correct decision interaction with effect size $f = .10$ ($\alpha = .05$, $1-\beta = .80$, two-tailed test based on a mixed ANOVA) after accounting for outliers who might translate verbal quantifiers into excessively high or low amounts. We obtained data from 154 participants after excluding 11 submissions that failed a check for reading attention (see below). 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 translated four verbal quantities into numerical ones (presented in random order): low % fat, low % minerals, high % fat, and high % minerals. Participants' translations were then used as the numerical quantifiers in the decision-making task. Table 5 shows the distribution of participants' translations, which were on average lower than in Experiment 1.

Table 5.

Numerical translations of verbal quantifiers of fat and minerals given in Experiment 2 (*N* = 154).

	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%
IQR	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, Sirota, & Bonnefon, 2019)

Decision-making task. Participants then completed the same GDA decision task as in Experiment 1. The only difference was that in the verbal condition, the quantifiers were either low or high, and in the numerical condition, the quantifiers were participants' personal numerical translations for low fat, low minerals, high fat, and high minerals. 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 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 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.

[Table 6 near here.]

Table 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)

We measured decision performance and response time for each of the 8 decision trials. 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).

Results

We hypothesised that, in line with more intuitive processing, verbal quantifiers would result in quicker response times, poorer decision performance, and greater use of context-based shortcuts. As pre-registered, we conducted multilevel analyses using random by-participant intercepts (the full effects model did not converge). As shown in Table 2, the speed of participants' decisions was not affected by format, $F(1, 1130) = 1.77$, $p = .184$, $M_{\text{diff}} = 0.01$, 95% CI [-0.06, 0.07]. However, participants' decision performance was lower with verbal than numerical quantifiers, $F(1, 1130) = 22.63$, $p < .001$, $M_{\text{diff}} = 18.7\%$, 95% CI [11.53, 25.93] (see Table 4)[1].

Participants used nutrient-based shortcuts in their decisions

As shown in Table 3, the proportion of errors made when the correct response conflicted with an intuitive response for a nutrient was on average 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 correct decision 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, $F(1, 1130) = 3.95$, $p = .047$, $M_{\text{diff}} = 11.9\%$, 95% CI [0.15, 23.74] but not when the

[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, judging performance against the average translations reduced performance from 62% to 38%. This indicates that participants performed better when they were judged based on what they believed the verbal quantifiers to mean, rather than what they tend to mean on average, however this performance was still worse than their performance with numerical quantifiers.

quantifier was numerical, $F(1, 1130) = 0.20$, $p = .652$, $M_{\text{diff}} = 2.7\%$, 95% CI [-8.96 14.31].

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.

General Discussion

In two experiments, we aimed to test the effect of quantifier format (verbal or numerical) on four correlates of processing style derived from dual-process theory: subjective effort, response time, decision performance, and use of context-based decision shortcuts (Evans & Stanovich, 2013; Kahneman, 2011). Based on the assumption that verbal quantifiers are intuitive and numerical quantifiers are analytical, we would expect the effect of format on these variables to converge. However, 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. We did not find significantly worse performance for verbal than numerical quantifiers in Experiment 1, but in Experiment

2, participants were more accurate with numerical than verbal quantifiers. Experiment 2 used a simpler design, with fewer levels for quantities and nutrients, and crucially, Experiment 2 accounted for the fact that each participant might interpret the verbal quantifiers differently from an average translation. These factors would have contributed to the ability to detect the performance effect. Finally, 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. Participants' decision patterns suggested that context played a more important role in guiding decisions with verbal than numerical quantifiers.

Testing predictions from the dual-process framework

Dual-process theory suggests that intuitive (vs. analytical) processes are automatic (vs. deliberate), quicker (vs. slower), less effortful (vs. more effortful), and introduce more decision errors (vs. less errors; Evans, 2008; Kahneman, 2011). We would thus expect that if numerical quantifiers fit the dual-process typology of being more analytical, people would make better decisions that took longer and required more effort as compared to decisions with verbal quantifiers.

We found that 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 would make better decisions when the same quantity is presented numerically than verbally—as long as the accuracy criteria correspond with their interpretations of the verbal quantifier.

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, Brown, Moehlis, Holmes, & Cohen, 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, Zakay, & Usher, 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 'default-interventionist' view that intuitive processing produces responses that are quick but often need to be corrected by the slower analytical system (Sloman, 1996). However, challenges to this view posit that intuitive responses need not always require correction (e.g., Hammond, 1988, Bago & De Neys, 2017, 2019; Plessner & Czenna, 2008). Under the right circumstances, for example, if the task structure matches the type of intuitions generated, then that intuition can be accurate (Hammond, 1988; Rusou et al., 2013). More recently, Bago & De Neys (2019) found that people who got logical reasoning tasks correct often already had a correct instinctive response to the question, challenging the assumption that participants cannot be accurately intuitive on what appears to be an analytical task. Based on this view, participants could be processing numerical quantifiers intuitively, but accurately.

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. In our experiments, we presented a task that participants could complete 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, participants would be tempted to pair the nutrient in each trial with a certain response (e.g., minerals–healthy). These assumptions about the context are argued to be intuitive (De Neys, 2017a). We observed more errors in trials where the correct decision conflicted with the automatic associations, and verbal formats tended to produce a greater proportion of errors (nearly 50% more) than numerical formats.

Dual-process theories provide a framework to understand the error patterns found in our research. Intuitive decisions are typically described as relying on mental shortcuts (Kahneman, 2011). The nutrient in our case presented a mental shortcut to the problem based on existing knowledge 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). However, when the shortcut then creates an intuitive response that is in conflict with the correct response—a situation that results in more errors if that correction is not made (Bago & De Neys, 2017). Suppressing existing associations about the nutrient required a cognitive decoupling of information typically associated with analytical processing (Evans & Stanovich, 2013). Participants' patterns of decision-making could thus support for the proposition that verbal formats lead to decision biases (Windschitl & Wells, 1996).

The observed patterns of effort and reaction time, however, do not support a dual-process classification of verbal quantifiers as intuitive and numerical quantifiers as analytical. People were not slower with numerical quantifiers, nor did they find them harder than verbal quantifiers. Therefore, it is possible that the numerical format did not prompt more analysis, but different types of intuitions that better suited the task. For instance, numerical quantifiers might have activated intuitions about quantity magnitudes (e.g., 70% is close to 100% or 20% is close to 0%—these intuitions are more likely to lead to a correct answer in the task). On the other hand, verbal quantifiers might have activated intuitions about the context (e.g., minerals are more likely to be healthy, or fat is more likely to be healthy—these intuitions are less likely to lead to a correct answer in the task). Processing models such as the parallel constraint satisfaction model 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, Hilbig, & Jekel, 2014). An intuitive response made earlier in the decision process might then accurately integrate even seemingly complex information (Glöckner & Betsch, 2008; Trippas, Thompson, & Handley, 2017). In the case of our decision task, even though the time taken to make the decision was similar for verbal and numerical quantifiers, the decision and error patterns could differ based on whether the earlier activations were for the healthy or unhealthy response.

Alternative explanations of the results

Rather than verbal quantifiers being more intuitive than numerical ones, there are other characteristics of verbal quantifiers that might account for our results. For instance, other researchers posit that verbal quantifiers produce different decisions from numerical ones because they are more vague (Budescu & Wallsten, 1995), or elicit different construals of a problem (Ülkümen, Fox, & Malle, 2016), or convey information about the focal points of the quantifier (Moxey & Sanford, 1986; Teigen & Brun, 2000). Here, we consider these explanations in turn.

A vagueness explanation would argue that people naturally make more errors with verbal quantifiers because they do not translate them consistently (Budescu & Wallsten, 1995). However, we believe that performance differences in our experiments were not due to variability in translations for two reasons. First, a participant who interpreted a verbal quantifier as less than the average translation in Experiment 1 (e.g., interpreting 'low' as less than 20%) would more accurately judge healthy combinations, but less accurately judge unhealthy ones (and vice versa if the interpretation was higher than average). This would result in a similar error rate to a participant who saw the 20% quantifier and made errors in either direction. Therefore, there should not be a clear performance advantage based on vagueness alone. Second, we reduced the interpretational vagueness of verbal quantifiers by introducing the translation procedure in Experiment 2, which accounted for participants' interpretations of verbal quantifiers in the different contexts (fat vs. minerals). Similar patterns in performance and error types were still observed with this procedure, showing that these effects were not merely due to interpretational variability.

Aside from inter-individual variability, one could also consider that the imprecision of verbal quantifiers makes them less suited to a calculation task.

Although we depicted previously consumed quantities using a pie chart so that there would be a different format from both the verbal and numerical quantifiers, the pie chart was still a pictorial representation of a precise number. The numerical quantifiers were thus arguably still more suited to the task, especially as they were point estimates. Participants in the verbal condition may thus have been more tempted to use contextual information to circumvent the effort involved in combining imprecise quantities. This explanation can complement some dual-process explanations, as it explains why people might undertake context-based shortcuts—defined in the context of our task as the 'intuitive response'. Future research might test this explanation directly by manipulating the vagueness between the two formats—keeping the vagueness constant (e.g., presenting numerical quantifiers as a range) or not.

A construal-based account would argue that people construe verbal quantifiers as more subjective and numerical quantifiers as more objective information (Ülkümen et al., 2016; Løhre & Teigen, 2015). Based on this account, we would predict that participants who were given objective numerical information would be more aligned with the quantitative task goal. In contrast, participants who were given subjective verbal quantifiers would construe it more as a recommendation instead of objective quantitative information. If participants took a verbal label like 'low minerals' as a recommendation, we would expect a tendency to take the full label into account (e.g., 'low minerals' is unhealthy rather than just 'minerals' is healthy). Instead, our data showed that both low and high minerals tended to be judged as healthy. We believe this indicates an unsuccessful attempt to combine the quantities due to the link between 'minerals' and 'healthy'.

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). 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. This explanation does not exclude the possibility that participants were 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.

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 indicated 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, Samsa, & Rimer, 2001). Hence, before drawing a firm conclusion from our results and assuming, for example, that numerical quantifiers are always 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 based on integrating quantities 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 decision scenarios. This might be through using different task structures where context-based answers typically lead to correct decisions: for example, deciding whether a combination of low fat and high minerals is healthy, vs. 10% fat and 60% minerals. This judgement invites an intuitive and correct answer for the verbal quantifier, but a more analytical process to get the right answer for the numerical. Different contexts could also test whether verbal and numerical quantifiers show the patterns observed in our studies (e.g., calculating whether expected payoffs in gambling tasks achieve an acceptable amount). 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.

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.

References

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. doi: 10.1523/JNEUROSCI.4343-05.2006

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. doi: 10.3389/fpsyg.2015.01088

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

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, 25*, 257-299. doi: 10.1080/13546783.2018.1507949

Berry, D. C. (2006). Verbal labels can triple perceived risk in clinical trials. *Therapeutic Innovation and Regulatory Science, 40*, 249-258. doi: 10.1177/009286150604000302

Berry, D. C., Knapp, P. R., & 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. doi: 10.1111/j.2042-7174.2002.tb00602.x

Berry, D. C., Raynor, D. K., & Knapp, P. R. (2003). Communicating risk of medication side effects: an empirical evaluation of EU recommended terminology. *Psychology, Health & Medicine, 8*, 251-263. doi: 10.1080/1354850031000135704

Betsch, T., & Glöckner, A. (2010). Intuition in judgment and decision making: Extensive thinking without effort. *Psychological Inquiry, 21*, 279-294. doi: 10.1080/1047840X.2010.517737

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (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. doi: 10.1037/0033-295X.113.4.700

Budescu, D. V., Por, H.-H., & Broomell, S. B. (2012). Effective communication of uncertainty in the IPCC reports. *Climatic Change, 113*, 181-200. doi: 10.1007/s10584-011-0330-3

Budescu, D. V., Por, H.-H., Broomell, S. B., & Smithson, M. (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change, 4*, 508-512. doi: 10.1038/NCLIMATE2194

Budescu, D. V., & Wallsten, T. S. (1985). Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes, 36*, 391-405. doi: 10.1016/0749-5978(85)90007-X

Budescu, D. V., & Wallsten, T. S. (1990). Dyadic decisions with numerical and verbal probabilities. *Organizational Behavior and Human Decision Processes, 48*, 240-263. doi: 10.1016/0749-5978(90)90031-4

Budescu, D. V., & Wallsten, T. S. (1995). Processing linguistic probabilities: General principles and empirical evidence. *Psychology of Learning and Motivation, 32*, 275-318. doi: 10.1016/S0079-7421(08)60313-8

Childers, T. L., & Viswanathan, M. (2000). Representation of numerical and verbal product information in consumer memory. *Journal of Business Research, 47*, 109-120. doi: 10.1016/S0148-2963(98)00055-1

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

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.

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353-386. doi: 10.1080/09541440244000328

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

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223-241. doi: 10.1177/1745691612460685

Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1055-1075. doi: 10.1037/0278-7393.34.5.1055

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account *Cognition, 133*, 641-666. doi: 10.1016/j.cognition.2014.08.017

González-Vallejo, C. C., Erev, I., & Wallsten, T. S. (1994). Do decision quality and preference order depend on whether probabilities are verbal or numerical? *The American Journal of Psychology, 107*, 157-172. doi: 10.2307/1423035

Hammond, K. R. (1980). The integration of research in judgement and decision theory. Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1988). Judgment and decision making in dynamic tasks. Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Jaffe-Katz, A., Budescu, D. V., & Wallsten, T. S. (1989). Timed magnitude comparisons of numerical and nonnumerical expressions of uncertainty *Memory & Cognition, 7*, 249-264. doi: 10.3758/BF03198463

Juanchich, M., & Sirota, M. (2019). Do people really prefer verbal probabilities? *Psychological Research*. doi: 10.1007/s00426-019-01207-0

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.

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. D. Gilovich, D. W. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York, NY: Cambridge University Press.

Knapp, P. R., Gardner, P. H., Carrigan, N., Raynor, D. K., & Woolf, E. (2009). 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. doi: 10.1348/135910708X375344

Knapp, P. R., Gardner, P. H., Raynor, D. K., 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 & Counseling, 79*, 267-271. doi: 10.1016/j.pec.2009.10.002

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

Liu, D., Juanchich, M., Sirota, M., & Orbell, S. (2019). People overestimate verbal quantities of nutrients on nutrition labels *Food Quality and Preference*. doi: 10.1016/j.foodqual.2019.103739

Løhre, E., & Teigen, K. H. (2015). There is a 60% probability, but I am 70% certain: communicative consequences of external and internal expressions of uncertainty. *Thinking and Reasoning, 22*, 369-396. doi: 10.1080/13546783.2015.1069758

MacLeod, A., & Pietravalle, S. (2017). Communicating risk: Variability of interpreting qualitative terms. *EPPO Bulletin, 47*, 57-68. doi: 10.1111/epp.12367

Malam, S., Clegg, S., Kirwan, S., & McGinigal, S. (2009). *Comprehension and use of UK nutrition signpost labelling schemes.* London, UK: BMRB.

Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Science, 22*, 280-293. doi: 10.1016/j.tics.2018.02.001

Millisecond Software. (2015). Inquisit 4. Retrieved from https://www.millisecond.com.

Millisecond Software. (2016). Inquisit 5. Retrieved from https://www.millisecond.com.

Moxey, L. M., & Sanford, A. J. (1986). Quantifiers and focus. *Journal of Semantics, 5*, 189-206.

Moxey, L. M., & Sanford, A. J. (1993). *Communicating quantities: A psychological perspective*. Hove: Lawrence Erlbaum Associates Ltd.

Nicolas, L., Marquilly, C., & O'Mahony, M. (2010). The 9-point hedonic scale: Are words and numbers compatible? *Food Quality & Preference, 21*, 1008-1015. doi: 10.1016/j.foodqual.2010.05.017

Olson, M. J., & Budescu, D. V. (1997). Patterns of preference for numerical and verbal probabilities. *Journal of Behavioral Decision Making, 19*, 117-131. doi: 10.1002/(SICI)1099-0771(199706)10:2<117::AID-BDM251>3.0.CO;2-7

Parsons, S., & Bynner, J. (1998). Influences on adult basic skills. Factors affecting the development of literacy and numeracy from birth to 37. London, UK.

Pennycook, G., De Neys, W., Evans, J. S. B. T., Stanovich, K. E., & Thompson, V. A. (2018). The mythical dual process typology. *Trends in Cognitive Sciences, 22*, 667-668. doi: 10.1016/j.tics.2018.04.008

Peters, E., Dieckmann, N. F., Västfjäll, D., Mertz, C. K., Slovic, P., & Hibbard, J. H. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied, 15*, 213-227. doi: 10.1037/a0016978

Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs, 26*, 741-748. doi: 10.1377/hlthaff.26.3.741

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.

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*, 549-556. doi: 10.1079/PHN2003552

Rusou, Z., Zakay, D., & Usher, M. (2013). Pitting intuitive and analytical thinking against each other: The case of transitivity *Psychonomic Bulletin & Review, 20*, 608-614. doi: 10.3758/s13423-013-0382-7

Schroyens, W., Schaeken, W., & Handley, S. J. (2003). In search of counter-examples: Deductive rationality in human reasoning *The Quarterly Journal of Experimental Psychology, 56A*, 1129-1145. doi: 10.1080/02724980245000043

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

Storcksdieck genannt Bonsmann, S., Férnández Celemín, L., Larrañaga, A., Egger, S., Wills, J. M., Hodgkins, C., et al. (2010). Penetration of nutrition information on food labels across the EU-27 plus Turkey. *European Journal of Clinical Nutrition, 64*, 1379-1385. doi: 10.1038/ejcn.2010.179

Teigen, K. H., & Brun, W. (1995). Yes, but it is uncertain: Direction and communicative intention of verbal probabilistic terms. *Acta Psychologica, 88*, 233-258. doi: 10.1016/0001-6918(93)E0071-9

Teigen, K. H., & 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. doi: 10.1002/1099-0771(200007/09)13:3<345::AID-BDM358>3.0.CO;2-U

Teigen, K. H., & Brun, W. (2003). Verbal probabilities: A question of frame? *Journal of Behavioral Decision Making, 16*, 53-72. doi: 10.1002/bdm.432

Teigen, K. H., Juanchich, M., & Filkukova, P. (2014). Verbal probabilities: An alternative approach. *Quarterly Journal of Experimental Psychology, 67*, 124-146. doi: 10.1080/17470218.2013.793731

Trippas, D., Thompson, V. A., & Handley, S. J. (2017). When fast logic meets slow belief: Evidence for a parallel-processing model of belief bias. *Memory & Cognition, 45*, 539-552. doi: 10.3758/s13421-016-0680-1

Ülkümen, G., Fox, C. R., & Malle, B. F. (2016). Two dimensions of subjective uncertainty: Clues from natural language. *Journal of Experimental Psychology: General, 145*, 1280-1297. doi: 10.1037/xge0000202

Viswanathan, M., & Childers, T. L. (1996). Processing of numerical and verbal product information. *Journal of Consumer Psychology, 5*, 359-385. doi: 10.1207/s15327663jcp0504_03

Viswanathan, M., & Childers, T. L. (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. doi: 10.1016/S1057-7408(08)80029-0

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*, 632-638. doi: 10.1177/1745691612463078

Wallsten, T. S., Budescu, D. V., 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. doi: 10.3758/BF03334162

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. doi: 10.1016/S0738-3991(00)00161-0

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*, 181-192. doi: 10.1037/0022-3514.60.2.181

Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied, 2*, 343-364. doi: 10.1037//1076-898X.2.4.343

Zimmer, A. C. (1983). Verbal vs. numerical processing of subjective probabilities. In R. W. Scholz (Ed.), *Advances in Psychology* (Vol. 16, pp. 159-182). London, UK: Elsevier.