VILNIUS UNIVERSITY
FACULTY OF PHILOLOGY

JULIJA KALVELYTĖ
Study programme "English Studies (Linguistics)"

NUMERAL INTERPRETATION: EXAMINING THE INTERPLAY BETWEEN LEFT-DIGIT BIAS, VALENCE, AND LINGUISTIC FEATURES

MASTER THESIS

Academic Supervisor – Assoc. Prof. Dr. Alexandre Cremers

Vilnius

2024

Acknowledgements

I would like to express my sincere gratitude to my advisor, Assoc. Prof. Dr. Alexandre Cremers, for his invaluable guidance and support throughout the writing process. His expertise, encouragement, and sound advice helped shape every part of this text, as well as every methodological choice that precludes it.

Likewise, I am eternally grateful to the faculty professors who generously provided much needed guidance and expertise at particularly difficult points of the process. Finally, I would like to thank all of the participants of the study for their time and effort.

The experiment carried out as part of the study was made possible by funding from the Research Council of Lithuania and European Social Fund under Measure 09.3.3-LMT-K-712.

Abstract

Originally described as a somewhat puzzling aspect of erroneous price judgements (e.g., judging $**4**.99 as closer to $**4**.00 than to **$5.**00), left-digit bias has since been observed across a wide range of domains, such as medicine, education, and competitive sports. Individually, these observations have all been interpreted as consequences of the imprecise mental mapping between abstract magnitudes (i.e., •••••) and their symbolic representations (i.e., 5), invariant across all aforementioned domains, dimensions, and scales. However, much less work has been done in order to reconcile this psychological perspective with the principles of pragmatics, governing the context-specific constraints of numeral meaning. In an effort to bridge this gap, the current thesis aims to establish how left-digit bias, valence, and selected linguistic features contribute to speakers' numeral interpretations. Using a combination of experimental and corpus data, aspects of left-digit bias are examined across various scales, sentential contexts, and languages. Based on data from 116 participants, the experimental portion reveals the degree of bias exhibited by English speakers across 30 different measurement scales and sentential contexts, as well as its lack of interaction with attribute-level affective valence. A context-sensitive model of numeral processing is proposed on the basis of these findings. The main assumption of the model is corroborated by corpus data from three different languages. However, further research is needed to establish the sensitivity of the bias to additional linguistic features. The implications of these findings for the relevant subfields of psychology and linguistics are discussed.

Keywords: left-digit bias, left-digit effect, numeral processing, round numerals, modified numerals, approximation, degree semantics

Contents

Introduction

What makes shoppers ‘penny wise and pound foolish’?, ask Thomas & Morwitz (2005) in their seminal paper on the phenomenon now known as *‘leading-’* or *‘left-digit bias’*. The answer, easy as it may be to infer from the name alone, may just as easily raise a few eyebrows: a single digit. Defined simply as “the tendency to focus on the left-most digit of a number and partially ignore other digits” (Englmaier et al., 2018, p. 1), this psychological phenomenon has often been cited as a major cause of biased purchasing decisions (Manning & Sprott, 2009). However, just as prices are not the single, most important kind of numerical information which shoppers encounter in their everyday lives, so too is the bias not unique to the domain of retail.

In fact, a strikingly similar phenomenon has been observed across a variety of other domains, including sports (Foellmi et al., 2016), medicine (Husain et al., 2021), and education. For example, Foellmi and colleagues reviewed extensive performance statistics from major sports leagues, analyzing massive datasets to identify patterns in player evaluations, which showed evidence of bias. Similarly, Husain and colleagues relied on electronic health records, drawing from a wide patient population to examine how small numerical differences influenced medical decisions. Likewise, Olsen (2013) collected data from educational institutions, analyzing grading records to reveal how “very small changes” in the leading digits of school-level grade averages could impact ordinary citizens’ opinions of public schools (p. 1). Be it an opinion, evaluation, or choice, private or professional, monetary or medical, the extant literature indicates that any sort of numerical judgment, seemingly situated on any scale, may be susceptible to the aforementioned bias, at least to some extent.

However, in spite of this apparent ubiquity, many theoretical accounts of this phenomenon seem to rely on different versions of the same basic explanation as the one cited above: enabled by biased attention, it leads to biased evaluation, and, consequently, biased decision. While this, of course, remains a perfectly sufficient and parsimonious explanation for the many cases in which the relevant number may have been encountered in the context of a a standard (i.e., decimal) scale or a well-organized, clearly-labeled public record of a specific metric, this is not always the case.

Although it may sound rather trivial, between heights, weights, times, distances, scores, and numerous other measures, speakers cope with a wide variety of different scales and dimensions on a daily basis (González et al., 2019). Furthermore, much of this information is couched in approximation, exaggeration, understatement, and scale-specific degrees of rounding. Sometimes due to uncertainty (i.e., opting for one's best estimate), sometimes to avoid judgment (i.e., deliberately understating the true number), and sometimes to elicit envy or praise (i.e., deliberately overstating). However, to what extent *LDB* contributes to these processes remains to be established.

With this in mind, the aim of the present thesis is to establish how left-digit bias, henceforth abbreviated to *LDB*, valence, and selected linguistic features contribute to speakers' numeral interpretations. In order to achieve this, the following four objectives have been set:

1. To establish how *LDB* affects speakers' numeral interpretations across various sentential contexts and measurement scales.
2. To determine how the affective valence (Russell 1980, 2003) associated with each target attribute moderates this effect.
3. To outline a rudimentary model of biased numeral processing which accounts for these findings.
4. To test the main theoretical assumption of the model against numeral approximator preferences suggested by corpus data.

The thesis is structured as follows: The first three chapters provide a brief overview of the relevant theoretical notions from the fields of linguistics, behavioral and cognitive psychology pertaining to the concept of *LDB*. Chapter four presents the experimental study, measuring *LDB* across various numerical magnitudes, scales, and valence levels. Subsequently, chapter five outlines a tentative model of biased number processing. Chapter six describes an exploratory corpus study testing the basic assumption of the model against data from three different corpora. This is followed by the discussion and conclusions. Subsequent sections provide the Lithuanian-language summary, data sources, references, and tools. Finally, all relevant supplementary materials can be found in the appendices.

1. Left-digit bias

As briefly mentioned in the Introduction, the general definition of *LDB* may seem rather simplistic and unlikely to invite extensive debate. However, there seems to be some modest debate regarding which digits should be treated as 'visible' and which as 'opaque', i.e., where the boundary between 'left' and 'right' digits may justifiably be drawn at different levels of magnitude. To this end, Lacetera, Pope, and Sydnor (2012) and Busse et al. (2013), suggest that the visible component of, for example, 25,623 miles on an odometer would likely be 20,000 miles, and the visible component of, for example, 122,367 miles on an odometer is 120,000 miles. Meanwhile, List et al. (2023) refer to Strulov-Shlain (2022), who proposes that in price judgment contexts, the dollar amounts of a product are to be interpreted as visible and the cent values, conversely, as opaque. The authors adopt this boundary in their own study. Notably, but perhaps unsurprisingly, these left-right boundaries seem to align quite well with the magnitude-dependent notion of rounding proposed by Krifka (2007) and further developed in more recent works, such as Solt (2015, 2017).

In the interest of a broader, linguistically-focused discussion, a modified definition of *LDB* will be adopted for the remainder of this thesis. As it stands, the basic definition suggests that on a decimal scale, the leftmost digit would correspond to a cognitively salient threshold. This is often, but not always, the case. For example, consider such scales as time (24 hours, 60 minutes), or length (in Imperial units). As can be seen from these, as well as many other examples, cognitively salient, intuitively clear categories do not always coincide with the decade or hundred boundaries. Therefore, *LDB* will be discussed in scale-variant terms, assuming that the relevant, bias-indusing threshold may be individually determined by the relevant context and scale, as opposed to necessarily aligned with decimal boundaries.

2. Theoretical models of number processing

Although there are many more (for a review, see Dowker, 2024), for the purposes of the present thesis two prevalent models of numeral processing will be briefly introduced: The Triple Code Model, proposed by Dehaene (1995), and the Encoding-Complex Model, introduced by McCloskey et al. (1985).

Broadly, Dehaene's model posits three distinct representational codes: the visual Arabic code, the verbal word frame, and the analog magnitude code, each facilitating different numerical tasks such as reading, speaking, and estimating quantities. This model emphasizes the interaction between these codes and neural correlates, thus providing a comprehensive neurocognitive perspective. In contrast, McCloskey's model, by comparison, would be generally considered more linguistically-oriented, focusing primarily on the symbolic and syntactic processing of numbers within a single representational system. McCloskey's model suggests a detailed mechanism for how numerical symbols are manipulated and understood through linguistic rules, thereby prioritizing the linguistic and syntactic aspects of numerical cognition. Consequently, while Dehaene's model integrates multiple cognitive and neural components, McCloskey's model offers a more language-centric view of numerical processing.

To illustrate the processing stages, consider the task of solving a simple arithmetic problem such as *7 + 5*. According to Dehaene's Triple Code Model, this task would involve the visual Arabic code to recognize and understand the numerical symbols '7' and '5.' The verbal word frame would then come into play as the problem is internally verbalized (*"seven plus five"*). Finally, the analog magnitude code would be engaged to approximate and calculate the sum. This interaction of visual, verbal, and analog processing highlights the model's comprehensive approach, integrating different cognitive pathways to produce the correct result of *'12.'* This multifaceted processing underscores the model's emphasis on the neurocognitive interplay between different representational systems.

Conversely, McCloskey's model would approach the same problem by focusing on the linguistic and symbolic manipulation of the numbers. The numbers *'7'* and *'5'* are processed through a unified representational system that encodes their symbolic form and syntactic structure. The model would entail parsing the arithmetic operation through a series of syntactic rules, essentially converting the symbols and operations into an internal linguistic form that can be manipulated according to learned arithmetic rules. For instance, the problem *'7 + 5'* would be translated into a syntactic format that facilitates step-by-step linguistic processing, ultimately producing the sum *'12.'* This process emphasizes the role of linguistic structures and symbolic manipulation, illustrating the model's focus on the syntactic and language-based aspects of numerical cognition. Hence, McCloskey's model highlights the linguistic intricacies of numerical processing, distinguishing it from Dehaene's broader neurocognitive framework.

How might more recent theoretical models of number processing explain left digit effects? A number of models have been proposed to explain the typical tendencies observed in number and price comparison tasks. For instance, the holistic model (Moeller et al., 2015) posits that numerical symbols are internally represented as approximate magnitudes along a mental number line, guiding responses based on the precision of this conversion process (Dehaene et al., 1990). Thomas and Morwitz (2005) based a portion of their original account of digit bias on this idea, suggesting that magnitude encoding for multi-digit numbers begins upon encountering the leftmost digit, distorting perceived magnitudes towards it. Conversely, Verguts and De Moor (2005) argued for a decomposed processing, where multi-digit numbers are broken down into individual digit representations for comparison based on place-value structure. Meanwhile, Nuerk and Willmes (2005) suggested a hybrid model incorporating both holistic and decomposed processing.

Similarly, Dotan and Dehaene's (2020) recent model suggests that digits are first quantified independently, weighted according to their roles, and then merged into a whole-number quantity representation that informs judgments. Overall, while all scholars seem to agree that the underlying mental process is inherently biased and imprecise, the exact origin and scope of this imprecision remain a subject of continuous theoretical debate and conceptual refinement.

3. Numeral processing factors

3.1. Cognitive factors

3.1.1. Attention

Overall, attention plays a crucial role in number processing, influencing various aspects of numerical cognition. A number of experimental studies have shown that attention can impact numerical processing by inducing spatial shifts of attention based on number magnitude and response side compatibility. In other words, smaller numbers have been shown to direct attention to the left and larger numbers to the right, reflecting an intuitive spatial-numerical association (de Hevia & Spelke, 2009).

These findings suggest that the mental number line, a spatial representation of numerical magnitude, is actively navigated by attentional shifts during number processing. In procedural terms, this means that when participants are asked to compare two numbers, their attention would be expected to shift to the reference number and then move towards the target number, with the speed and direction of these shifts influenced by the numerical distance (i.e., magnitude difference) from the reference number.

3.1.2. Memory

Furthermore, numeral processing involves a complex interaction between working memory and long-term memory systems. For instance, the retrieval of arithmetic facts from long-term memory and the manipulation of numerical information in working memory are interdependent processes that collectively facilitate efficient number processing. Additionally, spatial-numerical associations are influenced by the activation of canonical number sequences in working memory, highlighting the role of short-term memory representations in number categorization tasks. Consequently, individuals' susceptibility to *LDB* may also depend on factors such as the ability to deliberately prioritize auditory or visual aspects of the numeral (Sokolova et al., 2020) during encoding (the latter being, seemingly, more susceptible to bias). However, the specifics of these processes will not be discussed at length as the relevant literature seems to depart from the specific topic of digit bias in one of two ways: either it is investigated, but only within the category of price, or relevant observations are to be gleaned from the standard digit span task, which has generally been assumed to elicit literal encoding (i.e., as a series of unrelated numbers, rather than as digits belonging to a single, large number). This, crucially for the present study, does not seem compatible with the notion of biased magnitude perception typically associated with digit bias.

3.2. Affective factors

3.2.1. VAD model

The VAD model of emotion, originally introduced by Russell in 1980 and further elaborated in 2003, provides a framework for understanding emotions based on three core dimensions: valence, arousal, and dominance. Valence refers to the positivity or negativity of an emotional experience, encompassing pleasant and unpleasant feelings. Arousal indicates the level of activation or energy associated with a particular emotion, ranging from low to high arousal states. Dominance reflects the degree of control or influence an emotion exerts over an individual's thoughts and behaviors. Among these dimensions, valence holds particular importance as it serves as a fundamental evaluative component of emotional experiences. It determines whether an emotion is perceived as enjoyable or aversive, shaping individuals' overall emotional responses to stimuli and influencing subsequent cognitive and behavioral reactions. As suggested by the theoretical link between *core affect* (Russell 1980; 2003, p. 147) and decision-making processes as well as some recent experimental findings (Zhong, 2022), valence may be one of the moderating factors of *LDB*.

3.2.2. Amplifying and mitigating effects of valence

Valence plays a significant role in amplifying our perceptions of language by imbuing words and expressions with emotional significance. Positive or negative valence can significantly impact how language is interpreted and experienced. For example, positive language may evoke feelings of happiness, enthusiasm, or satisfaction, enhancing our overall perception of the message conveyed. Conversely, negative language can trigger emotions such as sadness, anger, or fear, leading to more critical or defensive interpretations of the relevant message.

Similarly, valence influences our perception of numbers by attaching emotional connotations to numerical information. Positive numbers, such as high scores or large sums of money, are often associated with feelings of success, abundance, or joy, leading to favorable perceptions. In contrast, negative numbers, such as debts or losses, evoke emotions of disappointment, frustration, or anxiety, shaping our attitudes and reactions towards numerical data. Thus, valence serves as a powerful lens through which we interpret and respond to both language and numbers in various contexts.

Similarly, valence can disrupt our perception of numbers by influencing how individuals interpret numerical information. Negative valence associated with numbers, such as financial losses or unfavorable statistics, can evoke feelings of anxiety, frustration, or hopelessness, leading to biased interpretations and decision-making. This emotional interference can impede our ability to objectively assess numerical data and may result in irrational or avoidant behaviors in response to perceived threats. Conversely, it may be used to nudge individuals towards certain behaviors or choices Schindler et al. (2023), by aligning the valence of the main attribute relevant to the choice and digit bias, to amplify their individual effects. Thus, valence exerts a powerful influence on both language and numerical perception, shaping our cognitive processes and behavioral responses in significant ways.

3.2.3. Confidence and expertise

Somewhat surprisingly, digit bias has been shown to distort even relatively subjective, personal assessments. For example, a study by Li and Qiu (2023) revealed that digit biases significantly affect college applicants' self-assessments and decisions. For instance, students perceive a significant difference between just-below-round test scores, such as 1399 versus 1400, despite the negligible actual difference in performance. Goodman, Gurantz, and Smith further demonstrated how these leading digit differences impact perceptions of academic achievement on an individual level. Interestingly, Patalano et al., (2022) obtain similar results even in laboratory conditions, by asking participants to rate the applications of hypothetical admissions applicants.

Similarly, Donnelly et al. (2021) showed that even in low-stakes situations, biased temporal categorization may influence productivity estimates and lead individuals to overextend their efforts simply because the relevant digits create the false perception of a time period spanning longer than it objectively does. This appears to align with earlier findings byTang (2020), who investigated the bias on the scale of prices and calendar days.

Somewhat surprisingly, even experts from various fields are not immune to such biased behavior. This has been most extensively documented with regards to professional traders (Kansara et al., 2020). However, using regression analysis Campbell (2023) established that judges, despite their professional expertise and the gravity of their decisions, are not immune to biased thinking influenced by *LDB*. This bias potentially contributes to disparities in incarceration rates before and after individuals reach the age of 20, reflecting how even legal professionals can be swayed by numerical thresholds. In the medical field, Coussens (2018) highlighted increased testing and diagnosis rates just after age milestones like 40, while Olenski et al. (2020) and Akosman et al. (2024) noted disparities in treatment beyond age thresholds such as 80. These studies suggest that doctors may reclassify patients mentally based on these thresholds, affecting their clinical decisions despite the lack of formal guidelines recommending such changes. Wang et al. (2021) and Melucci et al. (2022) demonstrated similar biases in elderly cancer treatment rates. These findings suggest that medical professionals, like judges, are susceptible to *LDB* despite their extensive training and experience. One theory for why this might be the case suggests that relevant contextual knowledge may be overshadowed by leading digits, which are widely regarded as meaningful despite their arbitrary nature (Fukuma et al., 2023). Similarly, Olenski and colleagues propose that experts might exhibit digit bias in situations where contextual knowledge might not be immediately available.

3.3. Linguistic factors

Additionally, numeral processing is influenced by a number of linguistic features (for an extensive review, see Bahnmueller et al., 2018, Ganayim et al., 2020), such as: magnitude, salience, roundness, inversion, and presentation mode (i.e., written versus spoken language perception). In the following sections, each of these factors will be briefly discussed in turn.

3.3.1. Magnitude

Magnitude, as important as it may be to pragmatic investigations of numerals, is conceptually rather straightforward. Simply put, as in the case with non-symbolic numbers, magnitude is the largesse of a particular number. The larger the number, the greater the magnitude. Why might this matter? As discussed by Krifka, Solt, and many of their fellow linguists, magnitude does not only denote the size of a number. By extension, it conveys to us how loosely or imprecisely it may be used without being objectively false. The recent study by Hesse and Benz provides an excellent illustration for this property: suppose that you hear about a recent gathering – a concert, a protest, or an similarly communal event. If you find out that about a thousand people came, how different could the actual number of attendees be for you not to feel misled by such an approximation? Likewise, if you hear that about 110 people got married today, how much could the actual number differ for the approximation to be true? These constraints are closely linked to the relative magnitude and roundness of the numeral in question. Finally, suppose that 1002 people got married. As can be seen from this and any number of analogous examples, a large number does not necessarily allow a large margin of error. With regards to approximation. both roundness and magnitude matter.

3.3.2. Roundness

Thus, it may be worth discussing the role of roundness. It is by no means novel, having already been reported as a potential cause of *LDB* by economists Bhattacharya and Jacobsen (2012) as well as experimentally examined in Rosch's (1975) work on cognitive reference points. The foundational assumption is that on numerical scales, round numbers act as cognitive reference points. This basic premise has since been extended and refined by other scholars, showing that roundness can be conceived in general as well as in scale-specific terms. These later perspectives posit that which numbers may be considered as round may vary depending on the structure and typical units of the relevant measurement scale, i.e., 60 minutes and 24 hours may both be perceived as round numbers depending on the relevant numerical scale. Conversely, what may be considered as round in an abstract context, may not be round within the confines of certain scales (i.e., 50 on a 100 typical test scale versus 50 on the chronological scale). This suggests the possibility of some interesting and unexpected variations with regards to digit bias and various measurement scales. In experimental contexts, the three-tier classification of roundness proposed by Jansen & Pollmann (2001) is typically adopted as a baseline, as has been done in the experimental study discussed in chapter 4 (for alternative examples, see Cummins, Sauerland & Solt, . (2012)).

Interestingly, a more complex alternative based on the concept of so-called *fluency* has been suggested by marketing researchers King and Janiszewski (2011). The authors argue that people tend to prefer numerical stimuli because their processing is less effortful or, in other words, more fluent. Although their account focuses on measuring and enhancing the appeal of brand names containing numerical symbols, they base their notion of fluency on the differential accessibility of various arithmetic facts. Subsequently, numbers involved in frequently and easily accessible arithmetic facts may be perceived as more fluent and potentially more appealing. In their terms, round numbers may be processed more fluently, and therefore preferred, over non-round numbers.

3.3.3. Salience

Salience refers to the prominence or distinctiveness of numerical values in a given context, impacting how readily they are processed and remembered. Highly salient numbers, such as culturally significant numbers (e.g., 7, 10, 100), are processed more efficiently due to their entrenched cognitive and cultural associations. These salient numbers often appear in educational settings, religious texts, and common idioms, making them more familiar and easier to manipulate mentally. On the other hand, numbers that lack cognitive salience, such as arbitrary or less frequently encountered values, require greater cognitive effort to process and are more prone to errors in both recognition and computation. This disparity is evident in both small and large numerical values, but it is particularly pronounced for larger, non-salient numbers that demand more mental resources. Additionally, the level of formality in communication affects cognitive salience; formal contexts, such as academic or legal settings, often necessitate precise and less familiar numerical values, whereas informal contexts allow for more flexible and familiar numbers.

3.3.4. Inversion

Inversion, a linguistic feature where the order of numerical components is reversed (e.g., ‘four and twenty’ instead of ‘twenty-four’), can complicate numerical processing, particularly for larger values and in contexts requiring precise calculations. This feature, found in languages such as German and French, can introduce additional cognitive steps in translating the spoken or written format into a mental representation suitable for arithmetic operations. For small numerical values, inversion may be less problematic due to frequent exposure and familiarity, but for larger values, it can significantly increase the cognitive load and potential for errors. Inversion also interacts with formality levels; formal contexts might mandate adherence to standard numerical conventions, thereby reducing inversion's impact, while informal contexts might see more variability and potential confusion.

3.3.5. Representational mode

The distinction between written and spoken modes also plays a crucial role in numerical processing. Written numbers are often more precise and unambiguous, facilitating exact calculations and reducing the risk of misinterpretation. This precision is particularly important for large numerical values and contexts with high expectations of precision, such as academic research or financial reporting. Spoken numbers, however, can introduce variability due to pronunciation, regional accents, and potential for mishearing, leading to greater cognitive demands for accurate processing. This difference is especially significant in large numerical values, where spoken formats may require repetition and clarification to ensure understanding. Additionally, formality levels affect how numbers are communicated; formal settings typically use written formats to ensure precision and clarity, whereas informal settings might rely more on spoken communication, where roundness and cognitive salience play larger roles in facilitating comprehension.

3.3.6. Reading direction

By the same token, consider reading directions. Although it might seem like there would be little common ground between reading and counting, reading has been linked to the mental number line. More specifically, the inherent orientation of the number line is generally assumed to be left-to-right (i.e., with the smallest numbers being situated farthest to the left). However, some research (Núñez, Cooperrider & Wassmann 2012; Núñez, 2021) suggest that this inherent orientation can be temporality or even permanently altered in speakers of languages with right-to-left oriented scripts. In other words, the orientation of the number line is closely associated with an individual's typical reading and writing direction. Additionally, experimental studies show that this orientation is also reflected in counting routines. While English speakers typically count by drawing out a horizontal line, Chinese speakers have been found to count vertically.

3.3.7. Digit order

Similarly, digit order inevitably affects how we encode and communicate multi-digit numbers. However, this influence may not be completely restricted to Arabic numerals. Although less explicitly, the order of digits in verbal numerals (tens-first versus units-first) appears to matter, at least to a certain extent. This is evidenced by differences in the performance on various numerical tasks (Göbel et al. 2014). Namely. studies with speakers of German or Dutch have revealed inversion-related performance differences in tasks such as arithmetic processing (Lonnemann & Yan, 2015), estimation (Savelkouls et al., 2020), transcoding (Lachelin et al., 2022), as well as in certain markers of numerical processing such as the compatibility effect for number words (Nuerk et al., 2005; Bahnmueller et al., 2019). Furthermore, recent accounts have even related the way we process number magnitudes to syntactic processing, positing a "novel syntactic effect" (Lozin & Pinhas 2022). Although most notable in children (Bugden, Park & Brannon, 2022), a number of these effects seem to persist into adulthood (for a review, see Klein et al., 2013).

4. Experimental study: left-digit bias across scales

4.1. Goal

The overall goal of the experiment was to gauge the extent of the bias across various measurement scales. This was achieved by computing three measures: the overall presence (or absence) of bias, expressed by deviation from the target values, the degree of bias, expressed by the mean deviation scores, and the interaction (or lack thereof) with affective valence.

4.2. Methods and data

4.2.1. Participants

151 participants (mean age = 36.5, SD = 10.0, range: [21, 68]) were recruited via *Amazon Mechanical Turk* (Buhrmester, Kwang & Gosling 2011; Paolacci, Chandler & Ipeirotis 2010). All participants were native speakers of English. Participants received $1.00 for their participation. To ensure that all participants understood the task format, written instructions and one practice item were provided. Additionally, each participant gave his or her informed consent before beginning the experiment and filled out a short questionnaire upon completion.

4.2.2. Setup

The experiment was carried out online via the platform *PCIbex* (Zehr & Schwarz 2018), utilizing a one-factor, five-level Latin Square design. The design was centered on five number levels. Specifically, each experimental item was associated with one number from each of the following levels and situated on the relevant measurement scale:

1) *threshold — a round number on the scale* (e.g., SPF 30)

2) *slightly below* (e.g., SPF 29)
3) *slightly above* (e.g., SPF 31)
4) *significantly below* (e.g., SPF 20)
5) *significantly above* (e.g., SPF 45)

Within these scales of five, the *significantly above* and *significantly below* numbers were used as controls. To mitigate any potential order effects, the order of items was randomized for each participant. Additionally, each measurement scale was classified by context-specific valence into one of three categories:

1) *positive* (e.g., sun protection factor)
2) *negative* (e.g., price)
3) *neutral* (e.g., musical tempo)

4.2.3. Task

Participants were asked to complete a modified version of the *Number Line Estimation* task (Siegler et al., 2011; Siegler & Opfer 2003; Thompson & Opfer, 2010), widely used in psychology and cognitive science to study mathematical proficiency (Sasanguie et al. 2013, Schneider et al. 2018, Simms et al. 2016) and numerical cognition (Hornung et al. 2014, Muldoon et al. 2013, Sella et al. 2017). In addition to formal arithmetic, this encompasses many broader numerical skills, such as magnitude estimation (e.g., identifying the magnitude of *1000*) , number comparison (e.g., determining the larger of *82* and *28*) and notation interpretation (e.g., recognizing *0.7* and *0.70* as different notations of the same magnitude). In the bounded version of this task, participants are asked to place numbers from a range predefined by the experimenter, one-by-one, on a line. The line denotes a numerical interval, typically 0-100. In the unbounded version of the task, a predefined increment (i.e., 2) is marked instead of the maximum value of the interval (i.e., 100). No other numbers or context clues are given. The positions chosen by participants are thought to indicate their estimation of the relative magnitude of every target number.

This basic format was modified by associating each group of target numbers, or item, with a short context paragraph. For each item, participants had to place the given target number on a bounded number line by using a slider. The slider was configured to log responses on a scale of 0-100, independently of the range shown to participants. This setup was selected to enable the measurement of the dependent variable termed *deviation,* formally defined as the deviation between the proper position of the target number with regards to the slider scale boundaries, and the actual position at which participants placed the slider.

4.2.4. Stimuli

The experimental stimuli consisted of 30 different scenarios corresponding to a variety of measurement scales. For the complete list, please refer to Appendix 1. To test a greater variety of scales as well as accommodate the experiment participants, all quantitative information was provided in US customary units. The context paragraphs were invented by the authors. The range of contexts was deliberately chosen to include all possible categories discussed in previous *LDB* literature, the only notable exception being the category of *age*, the relevant studies of which appeared to be limited to bias in patient age (Brant et. al. 2022) and organ donor age (Jacobson et al. 2022). As constructing a full set of true minimal pairs which would be sufficiently natural, grammatical, and scale-appropriate was not feasible, the following criteria were used during sentence construction:

1) *no syntactically-ambiguous structures*
2) *no additional numerals beyond the target*
3) *no modal expressions*
4) *no past or future tense verbs (except where this would compromise aspects of naturalness or grammaticality)*
5) *no stereotypical or potentially-offensive content (e.g., woman on a restrictive diet)*

Likewise, the target numbers and scalar bounds were chosen based on the following criteria:

6) *absence of obvious ratio (e.g., 50 on a scale of 0-100)*
7) *scale-specific roundness (e.g., 60 minutes on a chronological scale)*
8) *wide range of roundness levels (collectively)*

Prior to experimental implementation, all items had been verified in terms of naturalness and grammaticality by a native speaker unaware of their purpose. An example of one experimental item can be seen in *Figure 1*.

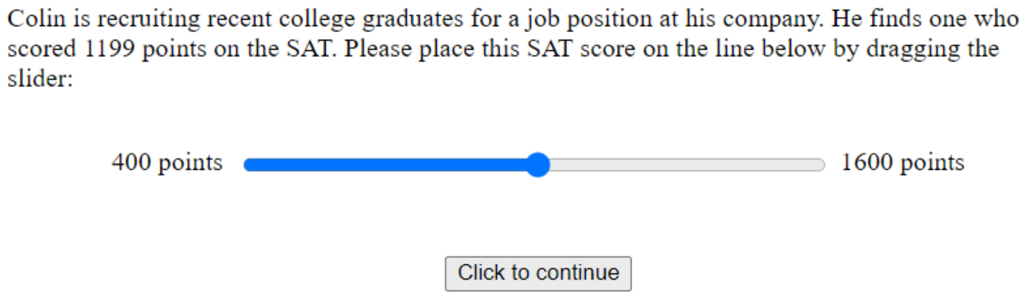

Figure 1.Example of an experimental item used in the Number Line Estimation task

4.3. Results

4.3.1. Pre-processing

In order to accurately interpret the experimental data, the results were processed, analyzed, and modeled in R (R Core Team 2022). First, all scale-specific values were normalized to percentages to enable comparisons with the participants' responses. Then, individual error rates were calculated based on the control items, in other words, the *significantly above* and *significantly below* category items. An individual error was formally defined as responses falling below the normalized threshold for the *clearly above* level items, and accordingly, above the threshold for the *clearly below* level items. The maximum error rate threshold was set to one standard deviation above the mean participant error rate. Based on the questionnaire data and the resulting error rate threshold of 38.5%, 35 participants were excluded from further analysis: 32 exceeded the threshold, two attempted to retake the experiment, and one exceeded the threshold and attempted to retake the experiment. The mean error rate for the remaining participants was 7.4%. All findings reported in the following sections are based on the remaining participants' responses.

4.3.2. Scalar variation

The global summary of *LDB* intensity across all tested scales can be seen in Figure 2 below.

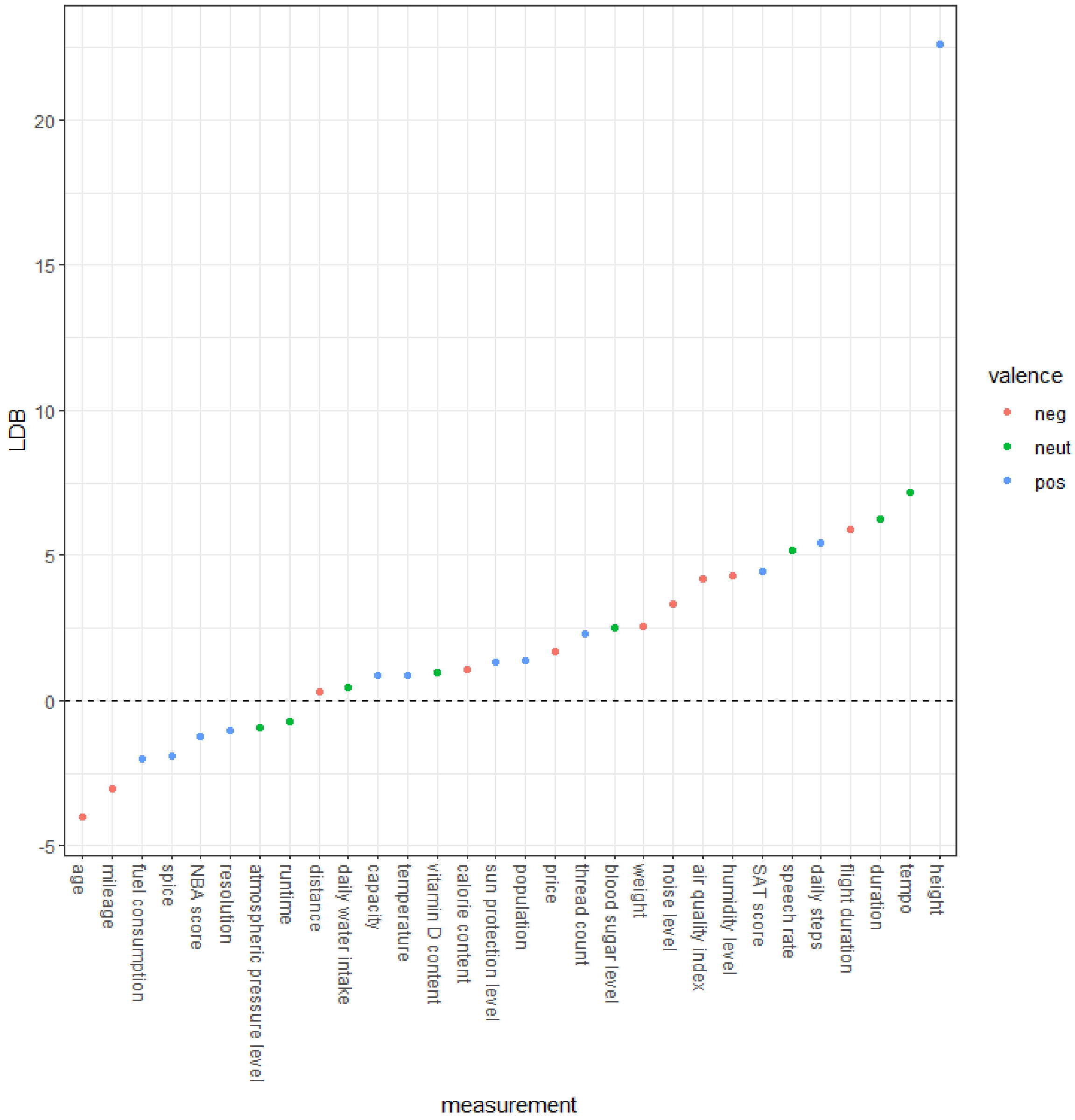

Figure 2.Summary of LDB across scales. LDB corresponding to the difference in mean deviation between the* threshold *and* slightly below *item placements, expressed in percentage points

As can be seen from the summary, responses associated with the *height* scale were simultaneously the most susceptible to positive deviation (mean = 0.23) as well as overall bias. Conversely, responses associated with the *age* scale were the most susceptible to negative deviation (mean = -0.04). No significant, consistent relationship between *LDB* and valence was identified, as evidenced by the relatively random distribution of *positive*, *negative*, and *neutral* valence responses.

4.3.3. Number level effects

After the initial pre-processing, mixed-effects linear models were fitted to the responses using the *lme4* package in R (Bates et al. 2015b). All models were fitted following the recommendations outlined in Barr et al. (2013) and Bates et al. (2015a) regarding the optimal random effects structure specifications. All estimates are given in percentage points.

Level	Mean	SE
Slightly above	-0.18	0.43
Slightly below	-1.35	0.47
Threshold	1.03	0.40
Clearly above	-1.29	0.35
Clearly below	2.24	0.46

Table 1.Mean and standard error values of each delineated number level

To test deviation with number level, the model was fitted to responses at the three relevant levels: *threshold, slightly above,* and *slightly below*. Responses associated with the *height* scale have been identified as outliers (see *Figure 2*) and excluded from any further analysis. However, note that they have not been excluded from any graph in the interest of accuracy and clarity. The model revealed a significant effect of number level ($\chi^2(2) = 9.25$, $p = 0.0098$). Although there was no significant deviation at the threshold level ($\beta = 0.8$, 95% CI [-1.0, 2.6]), deviations were significantly more negative at the *slightly below* level ($\beta = -1.7$, 95% CI [-2.8, -0.6]). By contrast, the deviations observed at the *slightly above* level were also more negative than at the *threshold* level, but not significantly so ($\beta = -1.1$, 95% CI [-2.4, 0.3]). Overall, the model possessed a moderate level of explanatory power (conditional $R^2 = 0.21$) Results of the first model have been summarized in Table 2 and Table 3 below.

Level	β	SE	T-value
(Intercept)	0.8	0.9	0.89
Slightly above	-1.1	0.7	-1.54
Slightly below	-1.7	0.6	-3.03

Table 2. Fixed effect estimates of the first mixed-effects model

Group	Level	Variance	SD	Correlation
Subject	**(Intercept)**	0.0009	0.03031	N/A
	Slightly below	0.0001099	0.0104	N/A
Item	**(Intercept)**	0.000478	0.0218	-0.62
	Slightly above	0.010124	0.10062	N/A

Table 3. Random effect estimates of the first mixed-effects model

4.3.4. **Valence effects**

To test the interaction of valence with position and deviation, a second model was fitted to the responses at the same three number levels, with the addition of valence as both a fixed and a random effect. To enable further statistical analysis, each of the three valence levels was converted to a corresponding numerical value (+1, 0, -1). The second model revealed no significant interaction between valence and number level (χ^2 (7) = 4.9, p = 0.67). (β = 0.5, 95% CI [-0.6, 0.7]), Overall, the model showed that a change in valence from negative to positive would lead to a 0.05 percentage point difference in deviation of *slightly below* level responses compared to *threshold* level responses. Overall, the model possessed a moderate level of explanatory power (conditional R^2 = 0.22). Results of the second model have been summarized in Table 4 and Table 5 below.

Level	β	SE	t value
(Intercept)	0.8	0.9	-0.90
Slightly above	-1.1	0.7	-1.55
Slightly below	-1.8	0.6	-3.22
Valence	-0.6	1.1	-0.55
Slightly above:Valence	-0.1	0.8	-0.18
Slightly below:Valence	0.1	0.7	0.15

Table 4. Fixed effect estimates of the second mixed-effects model

Group	Level	Variance	SD	Correlation
Subject	**(Intercept)**	0.000918	0.03030	N/A
	Slightly below: Valence	0.000416	0.02041	0.31
	Threshold:Valence	0.000332	0.0182	0.23; 0.57
Item	**(Intercept)**	0.001847	0.04298	-0.63
	Slightly above	0.0005337	0.02310	N/A

Table 5. Random effect estimates of the second mixed-effects model

4.4. Discussion

4.4.1. Main findings

In sum, these results illustrate the intricate dynamics between *LDB*, numeral properties, and affective valence. Notably, the bias is not limited to any particular measurement scale. Responses associated with the height scale exhibited the highest susceptibility to positive deviation (mean = 0.23), suggesting participants were more likely to overestimate these values, while those linked to the age scale showed a propensity for negative deviation (mean = -0.04), indicating the opposite tendency of minor underestimation. Importantly, the bias extends beyond standard scales. Contrary to our prediction, no consistent relationship between *LDB* and valence was identified, indicating that valence is not the determining factor of the bias. Furthermore, an examination of numerical level and deviation patterns revealed a systematic tendency to underestimate values slightly below the threshold, emphasizing the nuanced contextual factors influencing judgment outcomes. Although no significant interaction between valence and number level was observed, transitioning from negative to positive valence led to a minor difference in deviation for responses slightly below the threshold level. Further investigation is required to determine how exactly arousal and dominance may impact the interaction between digit bias and valence.

4.4.2. Comparison with previous results

As expected, the results reveal a minor but consistent *LDB* across all scales and contexts examined in earlier studies. However, given the wide range of scientific (i.e., determining the relative processing fluency of number formats) and practical (i.e., boosting sales) interests spanning the relevant body of literature, the possibilities of direct and detailed comparison may be limited and far-flung. Nonetheless, the following trends can be observed:

Although this factor was not central to the present study, it may be worth mentioning that the bias emerges rather early and persists into adulthood. Naturally, it is prevalent in adult magnitude estimations (Patalano et al., 2023). However, as demonstrated by Lai and colleagues (2018), it is already present in children as young as seven years old. In a study which compared the influence of leading digits on NLE performance between children (ages 7-11) and adults (ages 18-22), the authors found that although the bias diminishes with age, participants of all ages and number knowledge levels experience it to some degree. Importantly, it may also affect adjacent numerical abilities, such as number categorisation, in less experienced groups such as children.

Similarly, (Williams et al., 2022) investigated age-related bias trends in children (ages 5-8) and adults. Only 8-year-olds exhibited a significant *LDB* ($t(20) = 2.67$, $p = .015$, $d = 0.58$), with no such effect observed in younger children (all ts < 1.5, all ps > .10). Significant differences emerged around the decade boundary value 90 for 8-year-olds ($p < .001$), with 8- and 7-year-olds showing significant high-low difference scores (ts > 4.57, ps < .001). Conversely, adults displayed an average PAE of 3.6%, with strong left digit effects observed across trial blocks ($M = 0.74$, $SD = 1.28$, $t(43) = 3.85$, $p < .001$, $d = 0.58$), particularly around decade boundaries (ts > 2.00, ps < .001, ds = 0.71 – 0.89). While significant differences were found for hundreds difference scores on both the 0-1000 and 1000-0 lines (ts > 7.41, ps < .001), fifties difference scores were only significant on the 1000-0 line ($t(75) = 1.99$, $p = .050$, $d = 0.23$). Notably, left digit effects were consistent across line orientations, with no significant differences observed ($t(74) = 0.84$, $p = .406$).

Moreover, no main effects or interactions related to line orientation, task order, or gender were found (Fs < 1.35, ps $> .121$), suggesting a consistent influence of leftmost digits on estimates regardless of spatial orientation, with overall accuracy error similar across orientations.

In a later study, Williams and colleagues (2023) explored the left digit effect across different number ranges. They found that hundreds difference scores were significantly greater than zero for both the 0-1000 line ($M = 21.74$, $SD = 22.57$, $t(75) = 8.40$, $p < .001$, $d = 0.96$) and the 1000-0 line ($M = 21.48$, $SD = 25.28$, $t(75) = 7.41$, $p < .001$, $d = 0.85$). However, fifties difference scores were not significantly different from zero for the 0-1000 line ($M = 2.83$, $SD = 15.08$, $t(75) = 1.64$, $p = .106$), though they were for the 1000-0 line ($M = 3.55$, $SD = 15.6$, $t(75) = 1.99$, $p = .050$, $d = 0.23$). Comparing the left digit effect between the two line orientations, no significant difference was found ($t(74) = 0.84$, $p = .406$). They also observed that hundreds difference scores were significantly greater than fifties difference scores overall ($F(1, 74) = 52.96$, $p < .001$, $\eta 2p = .426$), irrespective of line orientation. Additionally, no main effects or interactions related to line orientation, task order, or gender were found (Fs < 1.35, ps $> .121$) on the left digit effect. These findings indicate that leftmost digits significantly influence individual estimates regardless of the spatial orientation of the number line, with overall accuracy error similar across orientations.

In terms of valence effects, the findings seem robust but somewhat mixed. Sabaghypour et al. (2023) investigated the interplay between valence and spatial biases using film clips as stimuli in a Numerical Line Estimation (NLE) task. They discovered that participants exposed to negative clips exhibited a significant leftward bias ($t(78) = 4.34$, $p < .0001$, $\eta 2 = .19$; $M = -1.27$, $SD = 2.61$), while those exposed to positive clips displayed a significant rightward bias ($t(79) = 5.28$, $p < .0001$, $\eta 2 = .26$; $M = 1.13$, $SD = 1.91$). Interestingly, regression analysis revealed that the leftward bias intensified with increasing number magnitude in the negative clip group ($F(1,23) = 5.34$, $p = .03$, $r2 = .19$), but this effect was not observed with positive clips ($F(1,23) = 1.83$, $p = .19$, $r2 = .07$).

In a similar study, White and Cohen (2022) investigated temporal perception among participants experiencing depression. They observed a significant main effect of valence ($F(2, 346) = 11.32$, $p < .001$, $\eta 2 = 0.06$), indicating differences in perceived accuracy error (PAE) among valence conditions. Post-hoc analyses revealed that the PAE for negative events significantly differed from that for neutral events but not from positive events. Conversely, the PAE for positive events significantly differed from that for neutral events, suggesting nuanced effects of valence on temporal distance perception in clinical populations, such as individuals with depression.

Expanding beyond symbolic estimation, Fabre et al. (2023) examined the influence of emotional states on non-symbolic estimation accuracy. They found that participants exhibited heightened accuracy in estimating concrete objects (e.g., cars) and reduced accuracy in estimating abstract objects (e.g., dots) when experiencing negative emotions compared to a neutral emotional state. This underscores the comprehensive impact of emotional states on both spatial biases and non-symbolic estimation accuracy.

Going another step further, Segal, Tzelgov and Algom (2024) carried out a study in which participants were asked to walk toward or away from a number, or say 'good' or 'bad' in response to a number. Overall, it took participants longer to say "good" than "bad" to small numbers, but it took them longer to say 'bad' than 'good' to larger numbers. The authors concluded that although individual participant outcomes could be accounted for by a spatial interpretation, the cumulative results are suggestive of the possibility of affective involvement in generating the effect.

As can be seen in *Figure 2*, the items within the *height* dimension yielded starkly different estimates. Furthermore, all test items were presented in a format that clearly delineated the unit and decimal digits. Although it may seem counter-intuitive, converging evidence suggests that the processes that underlie number reading may differ substantially from word reading, both in terms of chunking (i.e., how information is subdivided) (Dotan, 2023), landing position (i.e., where the eyes look first), and overall fixation patterns (i.e., how the eyes move) (de Chambrier et al., 2023). To this end, it might be worth briefly mentioning the mixed effects of visual separators with respect to written numeral processing. Namely, adding separators seemed to facilitate recall and comparison for healthy adults (M_{comma}=.42, $M_{nocomma}$=.57; $F(1,176) = 8.34$, $p= .006$) (Coulter et al., 2012, p. 401). However, this modification had no significant effect on number reading performance of dyslexic subjects (χ^2 =.79, p =.38) (Friedmann et al., 2010, p. 998). In the latter case, only the addition of so-called syntactic separators (i.e., ones or zeros) produced an effect.

4.4.3. Limitations

However, several limitations should be taken into account when interpreting the aforementioned findings. Firstly, the stimuli spanned a relatively limited range of numbers on each respective scale, restricting the generalizability of the findings. Additionally, the absence of filler items may have introduced some degree of so-called *demand effects* (Eckerd et al., 2021). The valence ratings were assigned manually. This, combined with the decision to focus on valence associated with the target attributes, rather than the sentences in which they were embedded (contrary to, for example, Siegel et al., 2020), may have introduced an additional level of subjectivity and variability into the results. However, as can be seen from the previous section, the relationship between valence and magnitude perception is rather unpredictable. Finally, the modified estimation task format, while uniquely suited to the purposes of this specific study, requires replication to fully substantiate its validity. Although multiple alternative task setups had been suggested based on some adjacent strands of the experimental literature (e.g., Savelkouls et al., 2020; Siegel et al., 2020; Wadhwa & Zhang, 2019), none have been deemed sufficiently methodologically rigorous or thematically relevant to pursue further.

4.4.4. Directions for future research

Thus, one obvious avenue for further research would be the replication of Experiment 1 with minor modifications. Likewise, the relationship between magnitude and valence may benefit from further linguistic investigation. However, this appears to be a relatively challenging subject to address meaningfully and exhaustively, as can be seen from a recent corpus study by (Bhatia et al., 2021). In addition to the subjectivity and variability discussed here, this study reveals an apparent plateau which readers experience when trying to process accounts of large-scale negative events. Lastly, a number of studies would seem to suggest that the expected level of precision for a particular context often relates to its perceived degree of (in-) formality. Thus, it may be reasonable to assume that in conversational contexts, digit bias may be moderated by formality.

5. Rudimentary model of biased numeral processing

Drawing upon the relevant notions of pragmatics and numerical cognition, the present section briefly outlines a model of biased numerical processing. Following prior works on the matter, the relevant modifying expressions will henceforth be collectively referred to as *approximators*[1].

Depending on the preceding discourse, the speaker's gaze and attention may be deliberately guided to a particular subsection of the number by a specific question or request from the listener (*"is it above or below 50 €?"*). This would circumvent the

[1] For alternative classifications, see Channell (2000), Eklund (2001), Nouwen (2010), Sauerland & Stateva (2011), Blok (2016), Baron (2022), among others

possibility of biased perception altogether. However, it seems reasonable to assume that, at least in most circumstances, the listener would not ask something quite as specific as *"is it a psychological price, ending in 95 or 99, or not?"* (Laurent & Vanhuele, 2023, p. 4). Additionally, it may be worth noting that, according to large-scale behavioral economics studies, *LDB* appears dissipate over larger values of the purchasing price (Anderson & Simester, 2003; Guéguen & Legoherel, 2004; Lin & Wang, 2017; Macé, 2012), However, corpus findings (Dehaene & Mehler, 1992; Jansen & Pollmann, 1996; Woodin et al., 2024) repeatedly demonstrate that smaller numbers are much more prevalent in communicative contexts. Therefore, no further comments will be made regarding these potential caveats.

By some accounts (Galaburda et al., 2002, p. 181), a direct transcoding pathway from the written Arabic number to the spoken numeral form may be possible, but only in cases where magnitude information is irrelevant to the task at hand (i.e., the number only needs to be read aloud). However, this idea has been actively debated and refined since its initial proposal. By the same token, a direct, non-linguistic pathway from Arabic numeral to abstract magnitude may be possible, as well as, susceptible to digit bias. Regardless of exact alternatives, it must be noted that the pathway described in the following paragraphs is certainly not the only one, and several candidate mechanisms of transcoding should be assumed by any alternative account.

Recall that in terms of basic numeral categories, a number may be round or not round. Suppose that a speaker needs to encode a particular number within the proximity of the round number N_0, which will simply be referred to as N, and that the goal is to achieve the most precise encoding with the least amount of processing effort possible. In order to account for the asymmetry implied by the relevant literature and corroborated by the experimental results, an additional distinction will be made between the numbers N_0 - ι (slightly below N_0), and $N_{0+\iota}$ (slightly above N_0). The magnitude of N will be referred to as N_M.

Assuming that N is read in a typical, unguided manner, its processing would proceed as follows: As the number is read, its digits would be scanned in a left-to-right sequence. Depending on the length (i.e., three-digits or longer), the number would be subdivided into chunks. Each three-digit chunk would then be processed in order. The identified digit string would then be converted to a precise number representation. At this stage, some magnitude information may be lost due to the speaker's inattention to the final digits of the string.

Subsequently, this exact representation would be encoded linguistically. The resulting intermediate representation may be exact or approximate, depending on the combination of pragmatic expectations and processing costs imposed onto the speaker by the communicative context. The former would be inferred from the conversational context. The latter would be informed by the speaker's world knowledge and inventory of available linguistic expressions. Notably, the former would include information about the most salient or relevant scale associated with the numeral *N*.

At this stage, the possibilities would be threefold: If *N* itself is a round number, it could simply be encoded as is without demanding any compromise between precision and processing ease. If *N* is not round, the remaining alternatives would be twofold: The number may still be encoded as is, achieving precision at the expense of ease, or with the help of an approximator, reducing precision but increasing the ease of processing.

More specifically, encoding may be facilitated by one of three types of approximators: sideward (e.g., *about* N_0), upward (e.g., *just above* N_0) or downward (e.g., *almost* N_0). Assume that conceptually, almost N_0 is more readily available than *just above* N_0, as the former can be reduced to a more basic conceptual form. From this it follows, that if a numeral has been encoded as *about N*, its magnitude should consequently be perceived as N_M. Conversely, if the numeral has been encoded as *almost* N_0, its magnitude should instead be perceived as significantly less than N_M. However, the cost and viability of each will vary depending on the number in question, namely: If the number is $N_{0\,-\,\iota}$, the upward approximator option will be unviable because $N_{0\,-\,\iota} < N_0$. Of the viable alternatives, *exactly N* will be the costliest in terms of processing effort but the most precise and the sideward alternative will be relatively less costly but the least precise. Thus, the downward approximator would be the most effective encoding option. Conversely, if the number is $N_0 + \iota$, the downward approximator will be unviable, because $N_{0\,+\,\iota} > N_0$. Of the remaining options, *exactly N* will again be very costly and *above* N_0 will be about as precise as *below* N_0 in the case of $N_{0\,-\,\iota}$, but relatively more costly.

As suggested by the relevant experimental findings, the degree of digit bias may be further exacerbated by negatively-valenced, and especially, highly-arousing content. Conversely, it may also be mitigated in contexts where *N* is presented in numeral form (i.e., *'seventy-nine'* instead of *'79'*) or in a language in which numeral morphology is inverted (i.e., *'neunundsiebzig'* instead of *'seventy-nine'*). This is attributed to a weaker conceptual association between magnitudes and numerals as compared with Arabic numbers (Greenstein & Velazquez, 2017) or more evenly distributed attention (Savelkouls et al., 2020). However, the extant literature would suggest that for this to occur, the speaker would have to deliberately focus on the verbal, rather than the visual, form of the numeral at the time of encoding.

After this stage, the selected linguistic expression would be translated into the corresponding abstract magnitude. This, in turn, may lead to biased magnitude perceptions and any equally biased judgements.

6. Exploratory corpus study: left-digit bias across languages

6.1. Goal

In order to , the experiment was followed up by an exploratory corpus study. The study is guided by the tentative assumption that the *just below* category expressions may correspond to a greater range of magnitude representations than those in the *just above* category. If this is indeed the case, a significantly greater number and variety of the former than the latter should emerge from the corpus data. Although there is no one-to-one mapping between numbers and inexact expressions such as,

6.2. Methods and data

6.2.1. Corpora

Overall, three corpora were consulted: the Corpus of Contemporary American English (Davies 2008–), the German Reference Corpus (Leibniz-Institut für Deutsche Sprache, 2022), and the Russian National Corpus (2003–). This selection was guided by several factors such as corpus size, text coverage, and varying degrees of numeral transparency associated with each of these languages (Moeller et al., 2015; Vasilyeva et al., 2015). While it may seem that these opaque aspects present challenges exclusively for young children and clinical populations, the findings discussed in the previous sections suggest that they could subtly affect numeric communication in ways which may merit further investigation.

6.2.2. Search and classification criteria

Following the principles outlined in Altenberg (2013), a survey of dictionaries and native speakers was conducted in order to accurately select and classify expressions corresponding to the two relevant concepts: *slightly below N* and *slightly above N*. Candidate expressions were collected by scanning the entries of four dictionaries (*Cambridge Dictionary of American English, Macmillan Thesaurus, Duden: Deutsches Universalwörterbuch, Novyi Slovar' Russkogo Yazyka [New Dictionary of Russian]*) for the following phrases: *by a small amount*, *almost*, *almost not*, *not quite*, *only just*. The initial selection was revised in accordance with the brief feedback obtained from native speakers of each language. Speakers remained naïve to the goal of the study. For a complete list of all selected approximators, please refer to *Appendix 2*.

Using the predefined list of relevant approximators and the built-in morphological annotation features, each corpus was searched for tokens matching the queries APPROXIMATOR + NUM (targeting e.g., *almost 100*) or NUM + APPROXIMATOR (targeting e.g., *100 something*). Where possible, several expressions were entered as alternants. Any relevant searches yielding a total of fewer than ten tokens were excluded from further analysis. The initial results were then reviewed manually. Although some degree of skepticism may be maintained with regards to the use of the built-in tagging system rather than a dedicated tagger, the publicly-available information on the accuracy rates of each corpus in question suggests an accuracy rate of 96-98% (see for example, Garside & Smith 1997 and Kuzmenko 2017). A summary of the overall distribution of the *slightly below* versus *slightly above* expressions across all three languages can be found at the end of section 6.3 below.

6.3. Results

6.3.1. American English

Beginning with the US English data, the following outliers were manually removed from the initial results: ordinal numerals, pronouns or line separators identified as Roman numerals, dates, chemical formulas, and addresses. Perhaps rather obviously, the latter tended to occur particularly frequently with approximators whose quantitative meaning has been derived from expressions of cardinal direction, such as *just South of* or *just North of*. This reduced the final results to an inventory of 50 different approximators across 117,225 tokens.

Of these, 31 were downward approximators, compared to 19 upward approximators. The most commonly-used approximators in the upward category appeared to be nearly, almost, and maybe. However, it should be noted that in the context of estimation, *maybe* tended to be quite ambiguous between its possible *slightly below*, *approximately*, and *not certainly* interpretations and thus may be overrepresented. It was followed by a substantially less frequent but much less ambiguous *close to*. The least common approximator in this category would seem to be *nigh*, which corresponded to only 13 results. This is not particularly surprising considering its strong associations with literary or old-fashioned styles and a high level of formality (see for example *Macmillan Thesaurus*). As a brief sociolinguistic aside, it may be worth noting that the equivalent query in the BNC yielded zero results, though this was most likely due to a difference in corpus sizes rather than actual regional preferences.

By contrast, the most frequent approximators in the slightly above category seemed to be *or so*, *just over*, and *barely*. However, the uniquely high frequency of the former expression may, to some extent, be due to the potential ambiguity between its slightly above and approximately interpretations. Interestingly, the three least common approximators were quite similar in terms of frequency, with *only just* yielding a mere 25 results, *and a bit* returning 26, and *just North of* adding up to 28 after the manual review. Collectively, expressions denoting numbers slightly below accounted for the overwhelming majority (85.4%) of the tokens extracted from the English corpus and 5.8 times as many as those denoting slightly above in terms of normalized frequency, reflecting a clear leftward bias. Importantly, the bias persisted even after excluding all of the potentially ambiguous approximators.

6.3.2. Russian

Turning to the Russian results, any non-numeric quantifiers (e.g., many) annotated as numerals were manually removed prior to further comparisons. Although the marking is grammatically accurate (see Zaliznyak 2003), the results themselves appeared to be semantically distinct from those most relevant to the overall goal of the thesis. Consequently, the final results consisted of 19 different approximators and a total of 15,614 tokens. Among these, the slightly below category encompassed 10 different approximators. The most frequent approximators in this category were *pochti* (lit. almost), *bez malogo* (lit. without a little) and *chut' ne* (lit. a bit not, nearly). Conversely, the least frequent approximator within this group was *bez mala* (lit. without a little), comprising a total of 13 results. This is not particularly unexpected as it is often characterized as the shorter, colloquial version of its more frequent counterpart, *bez malogo* (see *New Dictionary of Russian*).

Accordingly, the slightly above category comprised the remaining 9 approximators. In terms of frequency, *s lishnim* (lit. with surplus, with excess), *s lishkom* (lit. with surplus, with excess) and *s nebol'shym* (lit. with not big) appeared to be the three most favorable alternatives. Meanwhile, despite its rather unambiguous meaning and relatively simple form, *chut' vyshe* (lit. a bit higher) emerged as the least common choice with a total of 19 results. Notably, its lower counterpart *chut' nizhe* (lit. a bit lower) had been excluded at an earlier stage due to an even lower overall frequency of just 6 results.

Overall, expressions of the slightly below category encompassed more than two thirds (67%) of the tokens extracted from the RNC and twice as many as those placed into the slightly above category, suggesting a somewhat less prevalent leftward bias.

6.3.3. German

Finally, consider the German data. Using the morphologically-annotated DeReKo-2010-II subcorpus and the list of all relevant expressions, the following results were obtained: an array of 20 different approximators and 269,802 tokens. Within it, 11 expressions denoted numbers slightly below the reference point. Among these, the most frequent ones appeared to be *fast* (lit. almost), *gegen* (lit. against), and *nahezu* (lit. nigh, near to). However, the frequency of *gegen* may be somewhat overstated as, although it may be used to denote approximate quantity (see Pankau 2018), it is also commonly used in reference to opponents or competitors in much the same way as its English equivalent, *N against N*. Conversely, the least frequently-occurring one seemed to be *kaum weniger als* (lit. barely less than). The reasons for this may be rather obvious: it is the longer, more complex, less idiomatic alternative to its more frequent counterparts. Furthermore, it is quite similar in form to the lower-bound *nicht weniger als* (lit. not less than). One minor yet curious exception relates to *ein bisschen mehr als* (lit. a bit more than) and *ein bisschen weniger als* (lit. a bit less than). In spite of their close similarity, the former occurred 10 times and the latter only once, leading to its exclusion.

Perhaps somewhat counterintuitively, *etwas mehr als* (lit. some more than) was the second most frequent expression in the slightly above category, surpassed only by the short and simple *kaum* (lit. barely) and followed by *gerade noch* (lit. even still, just now). However, it should be clarified that the frequency associated with expressions with the adjectives *gerade* and *eben*, roughly equivalent to just (see for example Duden), may arise from one of two of its meaning components: quantity and recency, which may coincide or contrast depending on the exact measurement scale and context of use.

Overall, expressions belonging to the slightly below category accounted for the large majority of the tokens extracted from DeReKo (95%) and 30 times as many as those placed in the slightly above category.

Corpus	**Languag e**	**Tagger**	**Size (words)**	**Upward (raw freq.)**	**Downward (raw freq.)**	**Upward (norm. freq. pmw)**	**Downward (norm. freq. pmw)**	**Downward/Upwar d approximator ratios**
COCA	US English	CLAWS	950000000	17,195	100,060	18.1	105.3263158	5.819133469
DeReKo (DeReKo-2	German	CONNEXO R	1505820000	9,029	269,802	5.996068587	179.1728095	29.88171448

010-II (TAGGED-C-öffentlich archive))								
RNC	Russian	MyStem	374449900	5,149	10,465	13.75083823	27.94766403	2.032433482

Table 6. Summary of approximator frequencies

6.4. Discussion

6.4.1. Main findings

The corpus findings offer some tentative support to the theory outlined in the previous chapter. Overall, they reveal a notable prevalence of downward approximators across all examined languages, indicating a tendency towards underestimation. This pattern suggests traces of a similar bias extending from individual perception to communicative choices, with expressions denoting slightly lower numbers outnumbering their scalar equivalents. Interestingly, the bias appears to persist even when excluding the somewhat more ambiguous approximators.

6.4.2. Limitations

The aforementioned findings may have been affected by several limitations. Naturally, such a broad, macroscopic overview does not allow for many meaningful scale-specific or context-dependent insights to be drawn. Although additional means of obtaining such insights, successfully employed in earlier investigations of numerical quantifier choice (e.g., Williams & Power, 2013; Deckert, 2017; Sanko & Iriguchi, 2022) and potential affective impact (Nguyen et al., 2014), had been proposed, none have been deemed sufficiently relevant or precise to pursue.

It must be mentioned that, as it is often the case with native informant data, the final range of reviewed approximators may have been partially biased or restricted by the specific informants' linguistic experiences. Furthermore, a portion of the corpus data may also be biased by the somewhat artificial editorial guidelines governing the production of news texts, insofar as speakers (i.e., writers affiliated with the relevant publications) may forgo their natural intuitions in favor of editorial guidelines, which typically follow dedicated number format and rounding conventions. However, while this may affect individual speakers' rounding patterns in the corresponding professional contexts, this is rather unlikely to have affected the general orientation of the approximation preferences reflected in the data.

6.4.3. Directions for future research

One potential direction would be to investigate genre-based expectations of precision, building on experimental studies that indicate people generally assume newspapers exaggerate and sensationalize the true magnitude of reported events. Another research avenue could investigate how the interpretation of large figures is influenced by the perceived trustworthiness of sources, examining whether trusted sources lead to different valence associations compared to less trusted ones. Additionally, approximation patterns in languages with approximative inversion and other non-lexical means of expression present an intriguing area of study, potentially revealing how linguistic structures affect the perception and emotional impact of numerical information.

General discussion

Overall, the study revealed several notable patterns in digit bias, showing its dependence on both digit and overall number magnitude. This bias also proved to be context-sensitive and scale-dependent, extending to non-standard scales such as days (i.e., 24 hours). As evidenced by corpus data, speaker choices exhibited a consistent bias towards underestimation, regardless of numeral modification.

These outcomes resonate with previous studies that linked digit bias to numerical magnitude. However, the current findings further suggest that digit bias transcends traditional scales and contexts, making it a more pervasive phenomenon. The absence of a consistent relationship between digit bias and valence, unlike earlier research that documented valence effects on numerical estimation, highlights areas needing further exploration.

The results suggest that numbers *slightly below* scale-specific thresholds are perceived as further away, while those *slightly above* are seen as closer, potentially due to mental rounding, which may take place during estimation tasks. This underestimation bias, in turn, seems to indicate that cognitive and communicative factors may drive preferences for lower numerical estimates. The deviation observed on the height-related scale could be attributed to its complex structure as well as, the somewhat ambiguous link between target values and valence in this particular context (i.e., while desirable, this attribute may not necessarily be perceived as distinctly positive).

Several limitations must be acknowledged. The number line task and valenced sentence prompts might not fully encompass the complexity of left-digit bias and its interaction with linguistic features. The context-sensitive and scale-dependent nature of digit bias suggests that varying tasks or stimuli could yield different outcomes. Additionally, the influence of valence on digit bias might require more arousing stimuli or alternative task designs to be clearly observable. Future research should address these aspects to elucidate the underlying mechanisms of digit bias further.

Conclusions

Drawing upon the relevant experimental studies, a modified version of the number line estimation task was devised. Using this setup, speakers' magnitude estimations were measured across a variety of scales and sentential contexts. As predicted, estimation data collected from 116 English speakers revealed minor yet consistent digit bias. Notably, the bias was present across all tested magnitudes, scales, and sentential contexts, including cases in which threshold relevance could only be inferred pragmatically. These findings suggest that *LDB* does affect speakers' numeral perceptions in a context-sensitive manner.

Informed by the prevalence of emotional factors throughout the relevant body of experimental psychology literature (Choi et al., 2014; Dijksterhuis & Nordgren, 2006; Mikels et al., 2011), affective valence was examined as a potential moderator of *LDB*. Employing the same experimental setup as in the previous stage, a number of contextually-relevant scalar values associated with three valence levels (*negative*, *positive*, *neutral*) were tested. The experimental findings showed no significant interaction between *LDB* and valence. More specifically, the shift from negative to positive valence corresponded to a marginal increase of 0.05 percentage points in deviation for values associated with the *slightly below* compared to the *threshold* number level.

In an effort to fully account for these findings, a rudimentary model of biased numerical processing was proposed. Throughout it, the possible transcoding pathways were outlined, the main points of information loss were identified, and the stages potentially enabling the integration of relevant conceptual and linguistic information were proposed. Additionally, a more explicit and intentional mechanism of approximate encoding was suggested by means of round or modified numerals.

As an additional verification measure, an exploratory study of three different corpora (*COCA, RNC, DeReKo*) was carried out. The results aligned well with the main theoretical assumption underpinning the model, demonstrating a reliable speaker preference for expressions corresponding to values falling just below, as opposed to just above, the reference number. Notably, although exact ratios varied, the general preference remained constant across all corpora. This suggests that the bias may be a genuine aspect of numeral processing, rather than simply an artifact of the model or experimental setup.

Summary

Iš pradžių apibrėžtas kaip veiksnys, lemiantis klaidingą kainų suvokimą (pvz., **$4**.99 kainos suvokimą kaip intuityviai artimesnės **$4.**00, nei **$5.**00) vėliau kairiojo skaitmens nuokrypis buvo pastebėtas ir su kitų sričių (pvz., medicinos, švietimo ar varžybinio sporto) sprendimams aktualios skaitinės informacijos suvokimo bei vertinimo procesuose. Šios nuokrypio apraiškos buvo traktuojamos kaip netikslios asociacijos tarp abstrakčių mastų (t. y. •••••) ir jų simbolinės raiškos priemonių (t. y. 5), pasekmės. Tačiau ši psichologijos tyrimuose išsakoma pozicija iki šiol niekaip nebuvo siejama su pragmatiniais principais, lemiančiais skaitvardžių interpretacijos skirtumus įvairiuose kontekstuose. Siekiant užpildyti šią spragą, šiame darbe tiriama, kaip kairiojo skaitmens nuokrypis, afektinė vertė ir tam tikri kalbiniai bruožai kartu lemia kalbėtojų skaitinių verčių interpretaciją. Remiantis eksperimentiniais ir tekstyniniais duomenimis, kairiojo skaitmens nuokrypio aspektai nagrinėjami įvairių matavimo skalių, sakinių, ir kalbų kontekstuose. Remiantis 116 dalyvių tyrimo rezultatais, eksperimentinėje darbo dalyje nustatomas nuokrypio stipris, būdingas anglų gimtakalbiams trisdešimtyje skirtingų matavimo skalių ir sakinių kontekstų, bei jo sąveikos su afektines vertes kintamuoju stygius. Rezultatų pagrindu siūlomas skaitvardžių apdorojimo modelis, jautrus nuokrypio, kalbiniems ir kontekstiniems faktoriams. Pagrindinę modelio prielaidą patvirtina trijų skirtingų kalbų tekstynuose atsispindinčios kiekio raiškos priemonių pasirinkimo tendencijos. Tačiau norint tinkamai įvertinti galimą nuokrypio sąryšį su kitais kalbiniais bruožais, reikalingi papildomi tyrimai. Aptariama šių išvadų reikšmė susijusioms psichologijos ir kalbotyros kryptims.

Data Sources

Davies, M. (2008–) The Corpus of Contemporary American English (COCA): 520 million words, 1990-present. http://corpus.byu.edu/coca/.

Duden Deutsches Universalwörterbuch. (2003). (5., űberarbeitete Auflage). Mannheim: Dudenverlag.

Efremova, T. F. (2000). *Novyj slovar' russkogo jazyka: tolkovo-slovoobrazovatel'nyj; svyše 136 000 slovarnych statej, okolo 250 000 semantičeskich edinic*. Izdat. Russkij Jazyk.

Landau, S. I. (Ed.). (1999). *Cambridge Dictionary of American English*. Cambridge University Press.

Leibniz-Institut für Deutsche Sprache (2022) Deutsches Referenzkorpus / Archiv der Korpora geschriebener Gegenwartssprache 2022-I (Release vom 08.03.2022). Mannheim: Leibniz-Institut für Deutsche Sprache. www.ids-mannheim.de/DeReKo.

Macmillan Thesaurus. Available at: https://www.macmillanthesaurus.com/us/ (Accessed: 10 May 2023).

The Russian National Corpus (RNC). (2003–). http://www.ruscorpora.ru.

References

Akosman, I., Shafi, K., Subramanian, T., Kazarian, G. S., Kaidi, A. C., Cunningham, M., Kim, H. J., & Lovecchio, F. (2024). Left-digit bias in surgical decision-making for lumbar spinal stenosis. *The Spine Journal*. https://doi.org/10.1016/j.spinee.2024.03.007

Bahnmueller, J., Nuerk, H. C., & Moeller, K. (2018). A taxonomy proposal for types of interactions of language and place-value processing in multi-digit numbers. *Frontiers in Psychology*, *9*, 1024.

Bahnmueller, J., Maier, C. A., Göbel, S. M., & Moeller, K. (2019). Direct evidence for linguistic influences in two-digit number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(6), 1142–1150. https://doi.org/10.1037/xlm0000642

Bhattacharya, U., Holden, C. W., & Jacobsen, S. (2012). Penny wise, dollar foolish: Buy–sell imbalances on and around round numbers. *Management Science*, *58*(2), 413-431.

Brant, A., Lewicki, P., Wu, X., Sze, C., Johnson, J. P., Basourakos, S. P & Shoag, J. E. (2022). Impact of Left-Digit Age Bias in the Treatment of Localized Prostate Cancer. *The Journal of Urology*, *208*(5), 997-1006.

Busse, M. R., Knittel, C. R., & Zettelmeyer, F. (2013). Are Consumers Myopic? Evidence from New and Used Car Purchases. *American Economic Review*, *103*(1), 220–256. https://doi.org/10.1257/aer.103.1.220

Bugden, S., Park, A. T., Mackey, A. P., & Brannon, E. M. (2021). The neural basis of number word processing in children and adults. *Developmental Cognitive Neuroscience*, *51*, 101011.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/17456916

Campbell, Z. (2023). Happy Birthday, Go to Jail: Age-based Left-digit Bias in Criminal Sentencing. Unpublished manuscript. Department of Economics, Columbia University.

Choi, J., Jessica Li, Y., & Samper, A. (2019). The Influence of Health Motivation and

Calorie Ending on Preferences for Indulgent Foods. *Journal of Consumer Research*, *46*(3), 606–619. https://doi.org/10.1093/jcr/ucz002

Coulter, K. S., Choi, P., & Monroe, K. B. (2012). Comma N' cents in pricing: The effects of auditory representation encoding on price magnitude perceptions. *Journal of Consumer Psychology*, *22*(3), 395–407. https://doi.org/10.1016/j.jcps.2011.11.005

Coussens, S. (2018). *Behaving Discretely: Heuristic Thinking in the Emergency Department* (SSRN Scholarly Paper 3743423). https://doi.org/10.2139/ssrn.3743423

Cummins, C., Sauerland, U., & Solt, S. (2012). Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy*, *35*, 135-169.

de Chambrier, A.-F., Pedrotti, M., Ruggeri, P., Dewi, J., Atzemian, M., Thevenot, C., Martinet, C., & Terrier, P. (2023). Reading numbers is harder than reading words: An eye-tracking study. *Acta Psychologica*, *237*, 103942. https://doi.org/10.1016/j.actpsy.2023.103942

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 626–641. https://doi.org/10.1037/0096-1523.16.3.626

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 314–326. https://doi.org/10.1037/0278-7393.21.2.314

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 626–641. https://doi.org/10.1037/0096-1523.16.3.626

De Hevia, M. D., & Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition*, *110*(2), 198-207.

Dijksterhuis, A., & Nordgren, L. F. (2006). A Theory of Unconscious Thought.

Perspectives on Psychological Science, *1*(2), 95–109. https://doi.org/10.1111/j.1745-6916.2006.00007.x

Dotan, D., & Dehaene, S. (2020). Parallel and serial processes in number-to-quantity conversion. *Cognition*, *204*. https://doi.org/10.1016/j.cognition.2020.104387

Dowker, A. D. (2024). *Neural bases underlying numerical cognition* (1st ed.). SAGE Publications.https://ora.ox.ac.uk/objects/uuid:93f8ef4c-21c8-484b-a423-5ac95cdb1737

Eckerd, S., Du Hadway, S., Bendoly, E., Carter, C. R., & Kaufmann, L. (2021). On making experimental design choices: Discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes. *Journal of Operations Management*, *67*(2), 261–275. https://doi.org/10.1002/joom.1128

Fabre, L., Melani, P., & Lemaire, P. (2023). How negative emotions affect young and older adults' numerosity estimation performance. *Quarterly Journal of Experimental Psychology*, *76*(5), 1098–1110. https://doi.org/10.1177/17470218221107766

Foellmi, R., Legge, S., & Schmid, L. (2016). Do Professionals Get It Right? Limited Attention and Risk‐taking Behaviour. *The Economic Journal*, *126*(592), 724-755.

Friedmann, N., Dotan, D., & Rahamim, E. (2010). Is the visual analyzer orthographic-specific? Reading words and numbers in letter position dyslexia. *Cortex*, *46*(8), 982–1004. https://doi.org/10.1016/j.cortex.2009.08.007

Galaburda, A. M., Kosslyn, S. M., & Christen, Y. (2002). *The Languages of the Brain*. Harvard University Press.

Ganayim, D., Ganayim, S., Dowker, A., & Olkun, S. (2020). Linguistic effects on the processing of two-digit numbers. *Open Journal of Modern Linguistics*, *10*.

Garside, R., & Smith, N. (1997). A hybrid grammatical tagger: CLAWS4. In R. Garside, G. Leech, and A. McEnery (eds) *Corpus Annotation: Linguistic Information from Computer Text Corpora*, 102–121. London: Longman.

Greenstein, M., & Velazquez, A. (2017). Not all Anchors Weigh Equally. *Experimental*

Psychology, *64*(6), 398–405. https://doi.org/10.1027/1618-3169/a000383

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological science*, *25*(3), 789-798.

Altenberg, B., & Aijmer, K. (2013). *Advances in Corpus-based Contrastive Linguistics*. 1–305.

González, A. S., Hoffman, P., & Crutch, S. (2019). Where words meet numbers: comprehension of measurement unit terms in posterior cortical atrophy. *Neuropsychologia*, *131*, 216-222.

Goodman, J., Gurantz, O., & Smith, J. (2020). Take two! SAT retaking and college enrollment gaps. *American Economic Journal: Economic Policy*, *12*(2), 115-158.

Husain, S. A., King, K. L., & Mohan, S. (2021). Left-Digit Bias and Deceased Donor Kidney Utilization. *Clinical Transplantation*, *35*(6), e14284.

Hornung, C., Schiltz, C., Brunner, M., & Martin, R. (2014). Predicting first-grade mathematics achievement: The contributions of domain-general cognitive abilities, nonverbal number sense, and early number competence. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00272

Jansen, C. J., & Pollmann, M. M. (2001). On round numbers: Pragmatic aspects of numerical expressions. *Journal of quantitative linguistics*, *8*(3), 187-201.

Jacobson, C. E., Brown, C. S., Sheetz, K. H., & Waits, S. A. (2022). Left digit bias in selection and acceptance of deceased donor organs. *The American Journal of Surgery*, *224*(4), 1104-1108.

Kansara, A. P., Sung, M. C., Ma, T., & Johnson, J. E. (2020). Towards a better understanding of the full impact of the left digit effect on individual trading behaviour: unearthing a trading profit effect. *The European Journal of Finance*, *26*(9), 874-891.

Kim, J., & Kim, M. (2022). How do calorie‐ending and price‐ending matter in the food consumption decision making?. *International Journal of Consumer Studies*, *46*(3), 731-743.

King, D., & Janiszewski, C. (2011). The sources and consequences of the fluent processing of numbers. *Journal of Marketing Research*, *48*(2), 327-341.

Klein, E., Bahnmueller, J., Mann, A., Pixner, S., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2013). Language influences on numerical development—Inversion effects on multi-digit number processing. *Frontiers in Psychology*, *4*, 480. https://doi.org/10.3389/fpsyg.2013.00480

Krifka, M. (2007). Approximate Interpretations of Number Words: A case for strategic communication. Cognitive Foundations of Interpretation, 111-126.

Kuzmenko, E. (2017). Morphological analysis for Russian: Integration and comparison of taggers. In *Analysis of Images, Social Networks and Texts: 5th International Conference, AIST 2016, Yekaterinburg, Russia, April 7-9, 2016, Revised Selected Papers 5* (pp. 162-171). Springer International Publishing.

Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, *102*(5), 2206-2236.

Lachelin, R., Rinsveld, A. van, Poncin, A., & Schiltz, C. (2022). Number transcoding in bilinguals—A transversal developmental study. *PLOS ONE*, *17*(8), e0273391. https://doi.org/10.1371/journal.pone.0273391

Lai, M., Zax, A., & Barth, H. (2018). Digit identity influences numerical estimation in children and adults. *Developmental Science*, *21*(5), e12657.

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child development*, *78*(6), 1723-1743.

Laurent, G., & Vanhuele, M. (2023). How Do Consumers Read and Encode a Price?. *Journal of Consumer Research*.

List, J., Muir, I., Pope, D., & Sun, G. (2023). Left-Digit Bias at Lyft. *Natural Field Experiments*, Article 00771. https://ideas.repec.org//p/feb/natura/00771.html

Lonnemann, J., & Yan, S. (2015). Does number word inversion affect arithmetic processes in adults?. *Trends in Neuroscience and Education*, *4*(1-2), 1-5.

Lozin, M., & Pinhas, M. (2022). Perception of multi-digit number magnitude is dominated by numerical syntax: A novel syntactic end effect.

Manning, K. C., & Sprott, D. E. (2009). Price endings, left-digit effects, and choice. Journal of Consumer Research, 36(2), 328-335.

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*(2), 171–196. https://doi.org/10.1016/0278-2626(85)90069-7

Melucci, A. D., Loria, A., Ramsdale, E., Temple, L. K., Fleming, F. J., & Aquina, C. T. (2022). An assessment of left-digit bias in the treatment of older patients with potentially curable rectal cancer. *Surgery*, *172*(3), 851-858.

Mikels, J. A., Maglio, S. J., Reed, A. E., & Kaplowitz, L. J. (2011). Should I go with my gut? Investigating the benefits of emotion-focused decision making. *Emotion*, *11*(4), 743–753. https://doi.org/10.1037/a0023986

Moeller, K., Shaki, S., Göbel, S. M., & Nuerk, H. C. (2015). Language influences number processing–a quadrilingual study. *Cognition*, *136*, 150-155.

Muldoon, F. H., & Kuhlmann, H. C. (2013). Numerical error in modeling of particle-accumulation structures in periodic free-surface flows. *Computers & Fluids*, *88*, 43–50. https://doi.org/10.1016/j.compfluid.2013.09.005

Nuerk, H.-C., & Willmes, K. (2005). On the magnitude representations of two-digit numbers. *Psychology Science*, *47*(1), 52–72.

Núñez, R. E. (2021). From quantical to numerical cognition: A crucial passage for understanding the nature of mathematics and its origins. In *Heterogeneous Contributions to Numerical Cognition* (pp. 1-23). Academic Press.

Núñez, R., Cooperrider, K., & Wassmann, J. (2012). Number concepts without number lines in an indigenous group of Papua New Guinea. *PLoS One*, *7*(4).

Olenski Andrew R., Zimerman André, Coussens Stephen, & Jena Anupam B. (2020). Behavioral Heuristics in Coronary-Artery Bypass Graft Surgery. New England Journal of Medicine, 382(8), 778–779. https://doi.org/10.1056/NEJMc1911289

Olsen, A. L. (2013). Leftmost-digit-bias in an enumerated public sector? An experiment on citizens' judgment of performance information. *Judgment and Decision Making*, *8*(3), 365-371.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411–419. https://doi.org/10.1017/S1930297500002205

Patalano, A. L., Kayton, K., & Barth, H. (2023). Modeling the left digit effect in adult number line estimation. *Cognition*, *230*, 105257.

Patalano, A. L., Williams, K., Weeks, G., Kayton, K., & Barth, H. (2022). The left digit effect in a complex judgment task: Evaluating hypothetical college applicants. *Journal of Behavioral Decision Making*, *35*(1), e2247.

Pollmann, T., & Jansen, C. (1996). The language user as an arithmetician. *Cognition*, *59*(2), 219-237.

Rosch, E. (1975). Cognitive reference points. *Cognitive psychology*, *7*(4), 532-547.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145.

Sabaghypour, S., Farkhondeh Tale Navi, F., Kulkova, E., Abaduz, P., Zirak, N., & Nazari, M. A. (2023). The dark and bright side of the numbers: How emotions influence mental number line accuracy and bias. *Cognition and Emotion*, *0*(0), 1–14. https://doi.org/10.1080/02699931.2023.2285834

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number–space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, *114*(3), 418–431. https://doi.org/10.1016/j.jecp.2012.10.012

Savelkouls, S., Williams, K., & Barth, H. (2020). Linguistic inversion and numerical estimation. *Journal of Numerical Cognition*, *6*(3), 263-274.

Schindler, R. M., Isaac, M. S., & Wang, R. J. H. (2022). Strategic use of just-below numbers in packaged-foods calorie information. *Marketing Letters*, 1-14.

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta‐analysis. *Child development*, 89(5), 1467-1484.

Segal, H., Tzelgov, J., & Algom, D. (2024). Walking to a number: Is there affective involvement in generating the SNARC effect in numerical cognition? *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1384818

Sella, F., Berteletti, I., Lucangeli, D., & Zorzi, M. (2017). Preschool children use space, rather than counting, to infer the numerical magnitude of digits: Evidence for a *spatial mapping principle*. *Cognition*, *158*, 56–67. https://doi.org/10.1016/j.cognition.2016.10.010

Siegel, A. L., Graup, R. S., & Castel, A. D. (2020). Emotion-enhanced binding of numerical information in younger and older adults. *Quarterly Journal of Experimental Psychology*, *73*(1), 134–145. https://doi.org/10.1177/1747021819867292

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological science*, *14*(3), 237-250.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296. https://doi.org/10.1016/j.cogpsych.2011.03.001

Simms, V., Clayton, S., Cragg, L., Gilmore, C., & Johnson, S. (2016). Explaining the relationship between number line estimation and mathematical achievement: The role of visuomotor integration and visuospatial skills. *Journal of Experimental Child Psychology*, *145*, 22–33. https://doi.org/10.1016/j.jecp.2015.12.004

Sokolova, T., & Thomas, M. (2014). Focusing on the left digit: an encoding or an estimation bias?. *ACR North American Advances*.

Sokolova, T., Seenivasan, S., & Thomas, M. (2020). The left-digit bias: when and why are consumers penny wise and pound foolish?. *Journal of Marketing Research*, *57*(4), 771-788.

Solt, S. (2015). Measurement scales in natural language. *Language and Linguistics Compass*, *9*(1), 14-32.

Solt, S. (2017). About 40 or so people: How and why we approximate. *Bericht über das Forschungsjahr*, 45-53.

Strulov-Shlain, A. (2023). More Than a Penny's Worth: Left-Digit Bias and Firm Pricing. *The Review of Economic Studies*, *90*(5), 2612–2645. https://doi.org/10.1093/restud/rdac082

Tang, J. C. Y., Leung, S. C., Ying, K. H. D., Li, W. M., & Chan, W. L. (2010). The

Left-digit Effect in Price and Date Perception. In *Hong Kong Psychological Society Annual Conference*.

Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, *128*(2), 237–251. https://doi.org/10.1016/j.cognition.2012.09.012

Thomas, M., & Morwitz, V. (2005). Penny wise and pound foolish: the left-digit effect in price cognition. *Journal of Consumer Research*, *32*(1), 54-64.

Verguts, T., & De Moor, W. (2005). Two-digit Comparison: Decomposed, Holistic, or Hybrid? *Experimental Psychology*, *52*(3), 195–200. https://doi.org/10.1027/1618-3169.52.3.195

Wadhwa, M., & Zhang, K. (2015). This Number Just Feels Right: The Impact of Roundedness of Price Numbers on Product Evaluations. *Journal of Consumer Research*, *41*(5), 1172–1185. https://doi.org/10.1086/678484

Wang, J.-W., & Lin, C.-H. (2023). Temporal distance and motivation matter: Effects of psychological distance and left-digit in price evaluation. *Current Psychology*, *42*(13), 11069–11078.

White, M. R., & Cohen, D. J. (2022). Depression induced quantity estimation bias manifests only under time constraints. *Journal of Numerical Cognition*, *8*(1), 183–201. https://doi.org/10.5964/jnc.6515

Williams, K., Zax, A., Patalano, A. L., & Barth, H. (2022). Left Digit Effects in Numerical Estimation across Development. *Journal of Cognition and Development*, *23*(2), 188–209.

Williams, K., Xing, C., Bradley, K., Barth, H., & Patalano, A. L. (2023). Potential Moderators of the Left Digit Effect in Numerical Estimation. *Journal of Numerical Cognition*, *9*(3), 433–451. https://doi.org/10.5964/jnc.10091

Woodin, G., Winter, B., Littlemore, J., Perlman, M., & Grieve, J. (2023). Large-scale patterns of number use in spoken and written English. *Corpus Linguistics and*

Linguistic Theory, (0).

Zaliznyak A.A. (2003). Grammaticheskii slovar' russkogo yazyka. Izd. 4., ispravlennoe i dopolnennoe [Grammatical Dictionary of the Russian language] Ess. 4., corrected and supplemented. Moscow, Russkie slovari.

Tools

Bates, D., Maechler, M., Bolker, B., & Walker S. (2015b). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Wickham et al., (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686, https://doi.org/10.21105/joss.01686

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832.

Appendix 1

Full list of experimental items (continued on the next page):

Item	Measurement	Unit	Threshold	Significantly below	Slightly below	Slightly above	Significantly above	Valence	Minimum	Maximum
1	price	dollar	15	12	14.99	15.01	18	neg	0	40
2	runtime	minute	90	60	89	91	120	neut	0	200
3	resolution	megapixel	20	12	19.9	20.1	30	pos	0	50
4	weight	pound	200	150	195	205	300	neg	0	500
5	daily water intake	fluid ounce	60	45	59	61	80	neut	0	100
6	SAT score	point	1200	1000	1199	1201	1400	pos	400	1600
7	distance	mile	10	6	9.9	10.1	15.5	neg	0	30
8	atmospheric pressure level	hectopascal	1000	975	999	1001	1025	neut	950	1050
9	sun protection level	factor	30	20	29	31	45	pos	0	80
10	humidity level	percent	40	30	39	41	70	neg	0	100
11	duration	day	30	15	29	31	45	neut	0	100
12	age	year	30	20	29	31	40	neg	0	100
13	fuel consumption	mile/gallon	30	15	29	31	40	pos	0	50
14	air quality index	point	50	30	49	51	80	neg	0	100
15	population	inhabitant	15,000	8,000	14,900	15,100	22,000	pos	0	40,000
16	mileage	mile	100,000	80,000	99,900	100,100	120,000	neg	0	250,000
17	thread count	thread/inch	400	250	399	401	700	pos	0	1000
18	calorie content	calorie	700	500	699	701	800	neg	0	1000
19	daily steps	step	10,000	8,000	9,900	10,100	12,000	pos	0	15,000
20	noise level	decibel	80	62	79	81	95	neg	0	100
21	height	inch	6'0"	5'7"	5'11"	6'1"	6'4"	pos	5'0"	7'0"
22	flight duration	hour	24	20	23	25	30	neg	0	36
23	spice	scoville	2,500	2,000	2,499	2,501	4,000	pos	0	6,000
24	speech rate	word/minute	150	120	149	151	175	neut	0	200
25	NBA score	point	110	90	109	111	130	pos	0	200
26	vitamin D content	IU	600	300	599	601	800	neut	0	1000
27	blood sugar level	mg/dL	140	110	139	141	180	neut	100	200
28	temperature	degree Farenheit	60	25	59	61	80	pos	0	100
29	tempo	beat/minute	120	90	119	121	150	neut	0	200
30	capacity	gallon	50	35	49	51	65	pos	0	80

Context Paragraph
Kate is browsing for a t-shirt. She finds one that costs *. Please place this price on the line below by dragging the slider:
Ian is picking out a movie to watch on the weekend. He finds one that runs for * minutes. Please place this runtime on the line below by dragging the slider:
Dorothy is browsing for a camera phone. She finds one that takes pictures at * megapixels. Please place this resolution on the line below by dragging the slider:
Brandon is weighing himself before breakfast. His weight is currently * pounds. Please place this weight on the line below by dragging the slider:
Kevin is checking his water intake for the day. His total is currently * fl oz. Please place this daily water intake on the line below by dragging the slider:
Colin is recruiting recent college graduates for a job position at his company. He finds one who scored * points on the SAT. Please place this SAT score on the line below by dragging the slider:
Nicky is looking for the closest grocery store to do some last-minute shopping. She finds one that is * miles away. Please place this distance on the line below by dragging the slider:
Edward is checking the weather app on his phone. He sees that the barometric pressure outside is currently * hPa. Please place this atmospheric pressure level on the line below by dragging the slider:
Daniel is browsing for sunscreen. He finds one that is *. Please place this sun protection level on the line below by dragging the slider:
Lilly is checking the weather app on her phone. She sees that the humidity outside is currently *%. Please place this humidity level on the line below by dragging the slider:
Robert is looking for a project to work on until he starts his new job. He finds one that takes * days. Please place this duration on the line below by dragging the slider:
Jodie is casting actors for an upcoming film. She finds one who is * years old. Please place this age on the line below by dragging the slider:
Alicia is browsing for a new car. She finds one that gets * miles/gallon. Please place this fuel consumption on the line below by dragging the slider:
Daphne is checking the air quality in her neighborhood. She finds that the current air quality index is *. Please place this air quality index on the line below by dragging the slider:
Susan is looking up the population of nearby towns. She finds one that has * inhabitants. Please place this population on the line below by dragging the slider:
Angelica is browsing for a used car. She finds one with * miles. Please place this mileage on the line below by dragging the slider:
Albert is browsing for a pair of sheets. He finds one that has * threads/inch. Please place this thread count on the line below by dragging the slider:
Anabella is browsing for a salad bowl. She finds one that has * calories. Please place this calorie content on the line below by dragging the slider:
Sam is tracking his daily steps on a fitness app. His total is currently * steps. Please place this daily step count on the line below by dragging the slider:
Beth is browsing for a pair of noise-cancelling headphones. She finds one that can be used at up to * decibels. Please place this noise level on the line below by dragging the slider:
Jane is recruiting athletes for the local sports team. She finds one who is *. Please place this height on the line below by dragging the slider:
Jenelle is booking a flight for an upcoming trip to Australia. She finds one with a layover that takes * hours. Please place this flight duration on the line below by dragging the slider:
Zelda is browsing for chilli peppers to cook some curry. She finds a variety that is * scovilles. Please place this spice level on the line below by dragging the slider:
Darrell is browsing for an audiobook app in the app store. He finds one that can read aloud at * words/minute. Please place this speech rate on the line below by dragging the slider:
Luke is looking for an NBA team to bet on this season. He finds one that scored * points in their last basketball game. Please place this team's score on the line below by dragging the slider:
Andrew is browsing for a vitamin D supplement to take during the winter. He finds one that contains * IUs. Please place this vitamin D content on the line below by dragging the slider:
Victor is reading the results of a blood test he recently took. His blood sugar was at * mg/dL. Please place this blood sugar level on the line below by dragging the slider:
Marsha is checking the weather app on her phone. She sees that the temperature outside is currently * °F. Please place this temperature on the line below by dragging the slider:
Tom is looking for a song to play in the background as he works. He finds one that is * BPM. Please place this tempo on the line below by dragging the slider:
Ellen is browsing for a bathtub. She finds one with a capacity of * gal. Please place this capacity on the line below by dragging the slider:

Appendix 2

Complete list of all selected approximators:

English		Russian		German	
all but	only just	едва	почти	gerade erst	fast
almost	merely	с небольшим	без пяти минут	gerade noch	gegen
approaching	just past	с лишком	без малого	gerade so	nahezu
as good as	just over	с хвостиком	без мала	soeben	beinahe
nearly	just North of	чуть более	почти что	bloß	annähernd
nigh	just above	чуть выше	чуть не	kaum	praktisch
not far from	in excess of	с лишним	чуть менее	ein bisschen mehr als	quasi
not quite	hardly	с чем-то	чуть меньше	etwas mehr als	geradezu
on the order of	barely	с копейками	неполные	kaum mehr als	nicht ganz
practically	and some		практически		etwas weniger als
pretty much	and change				kaum weniger als
very nearly	and a bit				
virtually	a little more than				
nearing	or so				
maybe	a bit more than				
near	odd				
just under	some				
just South of	something				
just short of	ish				
just about					
in the vicinity of					
in the neighborhood of					
in effect					
essentially					
close to					
a little under					
a bit less than					
a little less than					
on the verge of					
on the edge of					
near to					