PREPRINT

Conceptual metaphor and graphical convention influence the interpretation of line graphs

Greg Woodin, Bodo Winter, and Lace Padilla, *Member, IEEE,*

Abstract—Many metaphors in language reflect conceptual metaphors that structure thought. In line with metaphorical expressions such as 'high number', experiments show that people associate larger numbers with upward space. Consistent with this metaphor, high numbers are conventionally depicted in high positions on the y-axis of line graphs. People also associate good and bad (emotional valence) with upward and downward locations, in line with metaphorical expressions such as 'uplifting' and 'down in the dumps'. Graphs depicting good quantities (e.g., vacation days) are consistent with graphical convention and the valence metaphor, because 'more' of the good quantity is represented by higher y-axis positions. In contrast, graphs depicting bad quantities (e.g., murders) are consistent with graphical convention, but not the valence metaphor, because more of the bad quantity is represented by higher (rather than lower) y-axis positions. We conducted two experiments (*N* = 300 per experiment) where participants answered questions about line graphs depicting good and bad quantities. For some graphs, we inverted the conventional axis ordering of numbers. Line graphs that aligned (vs misaligned) with valence metaphors (up = good) were easier to interpret, but this beneficial effect did not outweigh the adverse effect of inverting the axis numbering. Line graphs depicting good (vs bad) quantities were easier to interpret, as were graphs that depicted quantity using the x-axis (vs y-axis). Our results suggest that conceptual metaphors matter for the interpretation of line graphs. However, designers of line graphs are warned against subverting graphical convention to align with conceptual metaphors.

Index Terms—Conceptual Metaphor Theory, More is Up, Mental Number Line, Cognition, Linguistics, Emotional Valence, Line Graph, Axis Reversal, Handedness, Empirical Evaluation.

1 Introduction

Many data visualizations tap into existing cognitive associations. For example, it is conventional for numbers plotted on the y-axis of graphs to increase in magnitude from the bottom to the top of this axis [1], [2]. Consistent with this graphical convention, a large body of scientific research demonstrates that people tend to associate upward vertical locations with larger numerical quantities [3], [4], [5]. For instance, when participants are asked to spontaneously generate random sequences of numbers, they produce larger numbers when moved upward in a body-lifting device than when moved downward [6]. This empirical finding suggests that the upward ordering of numbers along the y-axis of graphs may not be arbitrary. Rather, this vertical orientation may have become established as conventional because it aligns with our vertical conceptualization of numerical quantities, making the resultant graphs easier to interpret [7], [8].

Cognitive associations where one concept (e.g., numerical quantity) is conceptualized in terms of another (e.g., vertical space) have been termed *conceptual metaphors* [9], [10], [11]. Conceptual metaphor theory is a framework based in cognitive science, linguistics, and psychology. The central tenet of conceptual metaphor theory is that metaphors in language are surface-level representations of deeper conceptual metaphors that structure the way we think about various concepts, including numerical quantity [3], [4], [12] and time [13], [14], [15]. For example, it is conventional for English speakers to describe numbers metaphorically as 'low', 'high', 'plummeting', or 'soaring' [11]. This linguistic pattern reflects the conceptual metaphor where larger quantities are associated with higher vertical positions. As well as language, the conventional design of data visualizations (e.g., the y-axis ordering of numbers) may reflect certain conceptual metaphors.

In the present study, we explored a subset of conceptual metaphors that may underlie graphical conventions in line graphs, focusing on metaphoric representations of numerical quantity and time. We investigated how the interpretation of line graphs is affected by the subversion of graphical conventions. We also explored whether line graph interpretation is influenced by conceptual metaphors of *emotional valence* [19], [20], [21], a term referring to whether concepts are perceived to be good or bad [22], [23], [24]. Emotional valence is relevant to the interpretation of data visualizations because, in many cases, the quantity depicted in a graph can be understood as good (e.g., vacation days) or bad (e.g., murders).

Our interest in emotional valence was inspired by a particular data visualization, widely discussed in the media [25], [26], blogs [27], [28], and scientific literature [1], [29], [30]. A graph designed in 2014 showed that gun deaths had increased since the introduction of Florida's Stand Your Ground law in 2005, which permitted state residents to use deadly force in self-defence (see Fig.1, A). The y-axis of this graph was inverted so that numbers increased from top to bottom rather than from bottom to top. While this inversion of death counts subverted graphical convention (the y-axis ordering of numbers), it aligned with experimental evidence showing that negatively valenced (bad) concepts (e.g., more deaths) are associated with lower vertical locations [19], [20], [21]. A previous study conducted by Pandey et al. [1] found that inverting the y-axis of graphs had an adverse effect on graph interpretability. However, this study did not consider the emotional valence of the quantities depicted in these graphs.

Motivated by the Stand Your Ground visualization (and other

- *G. Woodin and B. Winter are with the Department of English Language and Linguistics, University of Birmingham, UK. Corresponding author: G. Woodin. E-mail: gawoodin@gmail.com*
- *L. Padilla is with the Spatial Perception, Applied Cognition & Education Lab, UC Merced.*

Manuscript received month day, date; revised month day, date.

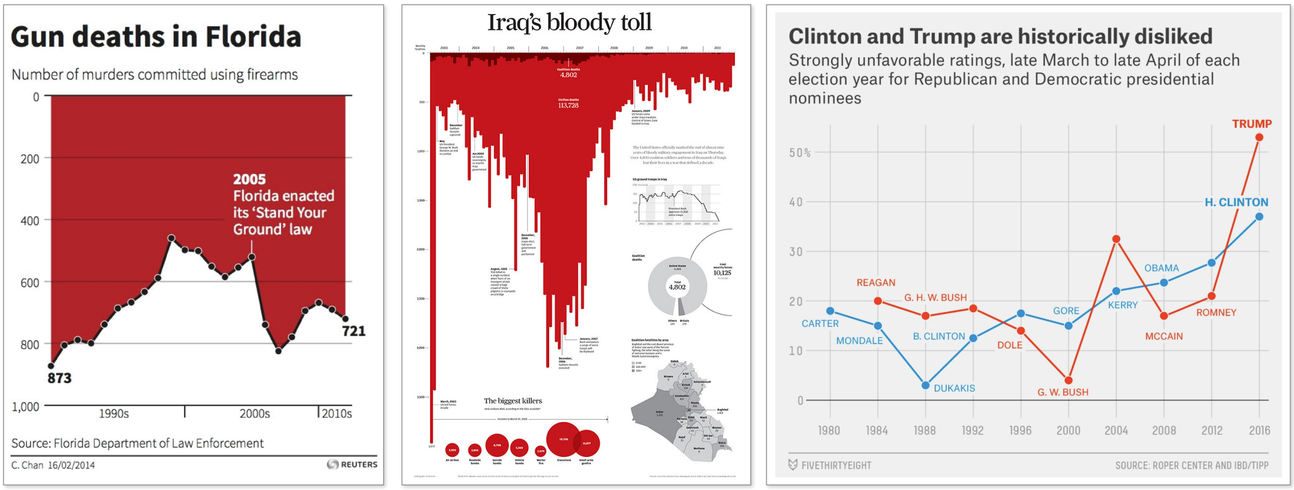


Fig. 1. Real-world data visualization examples that depict 'bad' (i.e., negatively valenced) quantities. A: Stand Your Ground visualization, designed by Christine Chan, published by Reuters [16]. B: Iraq's Bloody Toll visualization, designed by Simon Scarr, published by South China Morning Post [17]. C: Clinton and Trump Are Historically Disliked visualization, designed by Harry Enten, published by FiveThirtyEight [18].

visualizations; see Fig. 1), we conducted a pair of experiments investigating the factors that may affect the interpretation of line graphs depicting time series. We determined whether conceptual metaphors of emotional valence affect the interpretation of line graphs, and, if so, whether the extent of this influence is sufficient to justify departing from graphical convention (i.e., by inverting the axes of graphs). In doing so, we offer recommendations for designers with regard to how to spatially organise line graphs. The key contributions of this work are to:

- Provide a summary of existing research into a subset of conceptual metaphors that are relevant to the data visualization community (numerical quantity, time, and emotional valence)
- Explore the effect that subverting graphical convention has on the interpretation of line graphs (preregistered, conceptual replication-extension of Pandey et al. [1])
- Investigate the influence of conceptual metaphors of emotional valence on the interpretation of line graphs
- Determine whether designing line graphs to align with conceptual metaphors of emotional valence is justified, given that subverting graphical convention may adversely affect graph interpretability [1]

In the following sections, we begin by providing an overview of research into certain conceptual metaphors and their potential relevance to the design of data visualizations. We then describe two experiments. The first experiment focused on conceptual metaphors of emotional valence relevant to the ordering of numbers along the y-axis. The second experiment broadened this scope to include emotional valence metaphors relevant to the ordering of numbers along the x-axis. Our findings provide evidence that, all else being equal, designing line graphs to align with conceptual metaphors of emotional valence does have a beneficial effect on graph interpretability. However, the beneficial effect of this valence alignment on graph interpretability does not outweigh the adverse effect of subverting graphical convention. Based on this result, we recommend that line graphs should be designed to align with conceptual metaphors of emotional valence only if doing so does not contradict other well-established graphical conventions.

2 Background

As argued by Parsons [7], data visualizations can be made easier to understand by designing them with human psychology in mind, particularly by using knowledge of conceptual metaphors to inform design choices. Research in cognitive science, linguistics, and psychology has produced a plethora of evidence for the effect of conceptual metaphors on perception and behaviour (e.g., [6], [20], [31]). Scholars have previously discussed conceptual metaphors in the context of visualization research [2], [32], [33]. However, Parsons [7] notes that these scholars typically have not used metaphor research to offer actionable suggestions for visualization design (but see [8] for an exception). In the present study, we addressed this lack with an empirical investigation of a subset of conceptual metaphors that may influence the interpretation of line graphs. We also explored the relationship between conceptual metaphor and graphical convention, and the implications of this relationship for the design of line graphs. The relevance of conceptual metaphors to the data visualization community in general can be understood through the lens of *cognitive fit theory*, which provides a framework for understanding how people interpret visualizations.

2.1 Cognitive Fit Theory

Cognitive fit theory describes the mental effort required to understand data visualizations [34]. This theory proposes that people interpret visualizations with a mental model of what they expect to see, which informs the strategy used to interpret these visualizations [30]. For instance, the mental model used to interpret line graphs may include the expectation that numbers on the y-axis will increase from bottom to top. The Stand Your Ground visualization depicted in Fig. 1 (A) violates this expectation, reversing the direction of the y-axis so the trend line resembles blood dripping down a wall. This violation of expectation may force graph users to perform a mental transformation of the visualization, re-ordering the numbers on the y-axis to match their mental model. Such mental transformations are likely to increase cognitive processing costs, making the graph more difficult to interpret. For some graphs, the user may even arrive at an incorrect conclusion about the trend that the visualization depicts [29], [35], [36]. When viewing the Stand Your Ground visualization, for example, the graph user may mistakenly believe that gun deaths decreased rather than increased following the introduction of the Stand Your Ground law. Mental models may be informed by both graphical conventions and conceptual metaphors [7]. The first category of conceptual metaphor we discuss in this paper is relevant to the representation of numerical quantity and emotional valence along the y-axis of graphs – *vertical metaphors*.

2.2 Vertical Metaphors: Quantity and Valence

Research shows that people tend to associate upward vertical locations with larger numerical quantities and downward vertical locations with smaller numerical quantities [3], [6], [37]. Participants respond more quickly to larger numbers with a response button positioned higher in space and to smaller numbers with a response button positioned lower in space [3]. When participants are asked to spontaneously generate random sequences of numbers, they produce larger numbers when looking upward rather than downward [4], and when being moved upward rather than downward in a body-lifting device [6]. Gaze position also tends to shift upward when people count aloud in an ascending sequence [37]. This vertical conceptualization of numerical quantities is evident in the linguistic metaphors that English speakers use to talk about them. For example, speakers often describe numbers as ‘low’ or ‘high’, and may refer to ‘plummeting shares’ and ‘soaring costs’ [11].

Prior work suggests that the metaphoric association of larger numerical quantities with upward locations may form the basis for the conventional ordering of numbers upward along the y-axis of graphs [7], [8]. The benefit of graphical conventions aligning with conceptual metaphors may be that the resultant graphs fit a pre-existing mental model for thinking about numerical quantity and other concepts, decreasing cognitive processing costs and making these graphs intuitive and easy to understand [7].

There are, however, instances in which different conceptual metaphors make competing recommendations for how graphs should be designed. For example, the Stand Your Ground visualization (see Fig. 1, A) subverts the metaphoric association between larger numerical quantities and upward locations that is enshrined in graphical convention. However, this subversive design choice aligns with the vertical metaphor of emotional valence where positively valenced (good) concepts (e.g., fewer deaths) are associated with upward locations, whereas negatively valenced (bad) concepts (e.g., more deaths) are associated with downward locations. Evidence for this vertical metaphor of emotional valence shows that positively valenced words (e.g., ‘pride’) are evaluated more quickly when they appear in a higher position on a computer screen. Conversely, negatively valenced words (e.g., ‘liar’) elicit faster responses when they appear in a lower screen position [20]. Furthermore, participants recollect more positively valenced autobiographical memories when moving marbles upward from a lower box to a higher box than when performing the reverse vertical action [21], [38]. This conceptual metaphor is evident in linguistic expressions such as ‘in high spirits’, ‘down in the dumps’, ‘uplifting’, and ‘downbeat’ [11]. Conceptual metaphors of emotional valence are potentially important for graph comprehension because many quantities depicted in graphs have strong emotional connotations, such as income (increase = good) and infection rates (increase = bad) [39]. The fact that quantities depicted in graphs are often emotionally connotative raises the question of whether designers should factor vertical valence metaphors into the design of their visualizations.

The principal aim of the experiments reported in this paper was to establish whether conceptual metaphors of emotional valence influence the interpretation of line graphs. If they do, we investigated whether the influence of emotional valence is strong enough to justify inverting the y-axis of graphs, a design choice that subverts the conventional vertical ordering of numbers. A study conducted by Pandey et al. [1] has already shown that reversing the quantity axis of graphs – such that larger numerical quantities are represented by downward locations – has an adverse effect on the accuracy of responses to these graphs. Our experiments partly functioned as a preregistered, conceptual replication of this previous investigation, while also considering the effect of emotional valence. We showed participants line graphs with or without an inverted axis and asked these participants what the trend (i.e., the change in the quantity depicted in the graph over time) showed. We took the accuracy of responses to this question as a proxy for the interpretability of the graphs. Errors were taken to indicate that the graph in question contradicted the mental model that participants used to interpret the graph, in line with cognitive fit theory. We also extended this line of questioning to other conceptual metaphors, specifically ones that are relevant to the horizontal axis – *horizontal metaphors*.

2.3 Horizontal Metaphors: Quantity and Valence

Numerical quantities are not only conceptualized using the vertical axis, but also using the horizontal axis. The Spatial-Numerical Association of Response Codes (SNARC) effect shows that smaller numbers elicit faster left-hand responses, whereas larger numbers elicit faster right-hand responses [40], [41], [42]. Similarly, in random number generation tasks, leftward eye movements can be used to reliably predict that participants will generate a smaller number than the previous number in the sequence [5]. In addition, participants asked to place numbers anywhere on a piece of paper tend to orient the numbers horizontally from left to right [43]. These results have been interpreted as evidence for the existence of a *mental number line* in which people imagine numbers as increasing in magnitude from left to right. This horizontal conceptualization of numerical quantity may be the reason why certain graphs, such as horizontal bar charts, conventionally label numbers from left to right across the x-axis [44].

Unlike vertical metaphors, horizontal associations are typically not evident in language (e.g., we do not describe numbers as being ‘lefter’ or ‘righter’ than one another) [4], [45] and so are not always described as conceptual metaphors. However, we treat them as such because they deal with the same kind of cognitive process as prototypical conceptual metaphors that do have linguistic representations [45], [46]. To investigate horizontal metaphors of numerical quantity, in the present study we used line graphs that plotted quantity along the x-axis. We examined whether reversing the conventional direction of this axis – so that the numbers increased in magnitude from right to left – affected graph interpretability.

Like numerical quantity, emotional valence may also be conceptualized horizontally. When participants are asked to sort different items into boxes positioned on the left and right of a piece of paper, right-handers tend to sort positively valenced (good) concepts into the right-side box and negatively valenced (bad) items into the left-side box. Conversely, left-handers tend to do the opposite [47], [48], [49]. This finding is indicative of a conceptual metaphor where the dominant side of the body (and therefore that side of space) is associated with positive valence and the non-dominant side with negative valence. Complicating matters, because most people are right-handed, an association of right-side space with positive valence is evident in cultural practices (e.g., raising one's right hand when taking an oath) and language (e.g., the 'correct' answer is the 'right' one). The representation of right-handed conceptual metaphors in cultural practices and language may result in even some left-handers learning the association between rightward space and positive valence [50], [51].

In the present study, we investigated whether horizontal metaphors of emotional valence matter for the interpretation of line graphs that plot quantity along the x-axis. If so, we explored whether this effect was relative to handedness, or whether it was absolute (i.e., where both left- and right-handers associate right-side space with positive valence). If we found an effect of horizontal valence metaphors on graph interpretability, we also wanted to know whether this effect was significant enough to justify reversing the conventional ordering of numbers along this axis (i.e., ordering numbers from right to left rather than left to right). Previous findings presented by Pandey et al. [1] led us to expect that subverting the conventional x-axis ordering of numbers would have an adverse effect on graph interpretability. However, Pandey et al. exclusively inverted the y-axis of graphs, so our additional focus on x-axis inversion constituted an extension of this previous study.

2.4 Time Metaphors

The graphs used in the present study were line graphs that depicted changes in quantity over time, based on those used by Pandey et al. [1]. In a time series graph, which is one of the most frequently used types of data visualization [52], time is usually depicted as passing from left to right across the x-axis – earlier times are to the left, whereas later times are to the right [53], [54], [55]. The spatial representation of time in line graphs is consistent with horizontal time metaphors. For example, the Spatial-Temporal Association of Response Codes (STEARC) effect indicates that Western participants tend to associate leftward space with earlier times and rightward space with later times. When participants are tasked with indicating whether the timing of a sound is earlier or later than expected based on the temporal regularity of preceding sounds, participants respond to earlier sounds more quickly with a left-side button and to later sounds more quickly with a right-side button [31], [56] (see also [57]).

There are also vertical metaphors associated with time, with limited evidence suggesting that time is conceptualized as passing upward through space. For instance, one study showed that participants responded more quickly to future-related words when they appeared higher on a computer screen, and vice versa for past-related words [58] (see also [59]). In the present study, our use of line graphs depicting time series allowed us to test whether inverting the time axis of these graphs adversely affected graph interpretability. Our decision to invert the time axes of graphs extended the methodology used by Pandey et al. [1], who exclusively inverted the quantity axis of graphs.

2.5 Quantity Mapping

Our experimental stimuli included graphs in which quantity was 'mapped' onto either the vertical or horizontal axis, which we refer to as the graph's *quantity mapping*. We were therefore interested to see which quantity mapping would elicit more accurate responses. Previous research suggests that numbers are more readily conceptualized using the horizontal axis [43], [60]. However, results reported by Fischer et al. [44] suggest that vertically oriented bar charts are easier to interpret than horizontally oriented ones. Because line graphs conventionally depict quantity using the vertical axis, we predicted that graphs observing this convention would be easier to interpret. We predicted that graphs that subverted this convention by depicting quantity using the horizontal axis would be more difficult to interpret. Our choice to map quantity onto either the y-axis or x-axis also allowed us to test both horizontal and vertical representations of time.

2.6 Negativity Bias

A final possibility explored in our experiments was that the emotional valence of the quantity depicted in a graph matters by itself, irrespective of whether the graph aligns with valence metaphors. This possibility is suggested by the fact that people are quicker to recognise pleasant and safe stimuli than unpleasant or threatening stimuli [61], [62], [63]. Researchers have argued that negatively valenced stimuli contain more information than positively valenced stimuli, requiring greater attention and cognitive processing. This phenomenon has been termed the *negativity bias* [64], [65]. Based on the fact that positively valenced stimuli are processed more quickly and easily than negatively valenced stimuli, graphs depicting positively valenced quantities (e.g., vacation days) may have a cognitive processing advantage over graphs depicting negatively valenced quantities (e.g., murders).

3 Empirical Evaluation

We now report two experiments in which participants were asked questions about line graphs depicting a change in quantity over time. Experiment 1 focused on vertical valence metaphors. Experiment 2 served as a replication of Experiment 1, but also collected data from an equal number of left- and right-handers, allowing us to test horizontal valence metaphors.

3.1 Experiment 1

In this experiment, we tested the following hypotheses:

H1. Graphs without inverted axes will be easier to interpret than graphs with an inverted axis (preregistered, conceptual replication-extension of Pandey et al. [1]).

H2. Graphs that map quantity information onto the y-axis will be easier to interpret than graphs that map quantity information onto the x-axis (preregistered, conceptual replication of Fischer et al. [44] with line graphs rather than bar charts).

H3. Graphs depicting a positively valenced quantity (vacation days) will be easier to interpret than graphs depicting a negatively valenced quantity (murders).[1]

H4. Graphs aligning with vertical valence metaphors will be easier to interpret than graphs that do not align with these metaphors (see Table 1).

It should be stated from the outset that the experiments we present in this paper cannot disentangle the relative effects of

1. Hypothesis 3 was not preregistered prior to Experiment 1 but we include it here for consistency with Experiment 2 and as an additional test of this hypothesis.

TABLE 1
HYPOTHESIS 4: VERTICAL VALENCE ALIGNMENT

Stimuli	**Valence Aligning (y-axis)**
non-inverted y-axis, positive valence	yes
non-inverted y-axis, negative valence	no
inverted y-axis, positive valence	no
inverted y-axis, negative valence	yes

graphical convention and conceptual metaphor where both make identical recommendations for how graphs should be designed. For example, both graphical convention and conceptual metaphor dictate that numbers should increase upward along the y-axis of graphs, and that time should pass from left to right along the x-axis. If we were to find that inverting these axes had an adverse effect on graph interpretability, we would not be able to say whether this adverse effect were due to graphical convention, conceptual metaphor, or both. Therefore, although previous research has led us (and others [2], [32], [33]) to predict that conceptual metaphors shaped the formation of graphical conventions relating to numerical quantity and time, the present investigation is not intended to test this prediction. The only conceptual metaphors we investigate directly in this paper relate to emotional valence.

3.1.1 Procedure

The study was implemented using survey software Qualtrics [66] and distributed via Amazon Mechanical Turk [67], a crowd-sourcing marketplace where workers receive payment for completing tasks virtually. Our experimental procedure was inspired by Pandey et al. [1], particularly the aspect of their study that concerned the inversion of graphical axes. Participants answered multiple-choice questions about line graphs depicting made-up trends over time in fictional cities. The names of these cities were randomly generated using an online town name generator (https://www.namegenerator2.com/town-name-generator.php), with the criterion that each name should be six characters long. For example, one graph depicted the number of murders in Torley between 2015 and 2018 (see Fig. 2, left). Each graph showed a trend for a different city. Participants were asked: 'What can you say about the situation in this city?' Participants then indicated whether they believed the situation to be improving, declining, or neither improving nor declining (three multiple-choice options). In the Torley graph (Fig. 2, left), the number of murders in this city increased over time, indicating that the situation was declining. In contrast, if the graph showed a rising trend for the average number of vacation days (see trial types in Fig. 2, middle), the correct answer would be that the situation in the city was improving. The correct answer was never that the situation was neither improving nor declining.

Of primary relevance to our hypotheses was whether participants responded correctly or incorrectly. A higher proportion of correct responses to a graph was taken to indicate that this graph was easier to interpret than a graph that elicited a lower proportion of correct responses. The latencies of participants' responses were also recorded and investigated exploratorily.

The experiment proceeded as follows. In accordance with approved IRB protocol, after giving consent, participants were shown the instructions to the study (available at https://osf.io/5acjs/). These instructions included an exemplar graph that depicted the number of cars over time in a fictional city called Murell. Participants then answered an attention check question, being asked to accurately recall a particular word in the instructions they had just been shown. Participants who responded incorrectly to the attention check question were disqualified from completing the remainder of the study. Participants then completed four main experimental trials (see Fig. 2, middle), after which they supplied their demographic information.

3.1.2 Stimuli

We used a mixed between- and within-participants design. As shown in Fig. 2 (right), the type of graph shown to each participant was varied in a 3×2 design according to the following two factors:

- **Axis Orientation**: whether the graph (i) does not have inverted axes (right, top row in Fig. 2), (ii) has an inverted y-axis (middle row), or (iii) has an inverted x-axis (bottom row);
- **Quantity Mapping**: whether the graph maps quantity information onto its (i) y-axis (left column) or (ii) x-axis (right column).

These manipulations were conducted between participants to prevent participants from being cued to focus on the graph's axes more than they would naturally. This 3×2 design resulted in there being six groups, which were run as separate experiments on Amazon Mechanical Turk. Participants self-selected which group to participate in, rather than there being a strict randomization procedure. Participants were unaware of the group they were in as each person was only able to view one experiment. Responses from each group were collected until responses in each group reached the criterion for our preregistered stopping rule (N = 50).

Within each group, each participant completed four trial types, as shown in Fig. 2 (middle). This meant that each participant contributed four data points to the final data set. To generate the four trials within each group, we manipulated the following two factors within participants in a 2×2 design:

- **Trend**: whether the quantity depicted in the graph was (i) rising (middle, top row in Fig. 2) or (ii) falling (bottom row) over time;
- **Valence**: whether the quantity depicted in the graph was (i) positively valenced (left column) or (ii) negatively valenced (right column).

We chose 'vacation days' as a positively valenced quantity and 'murders' as a negatively valenced quantity. These two concepts were selected from Warriner et al.'s [68] list of 13,195 English words normed for emotional valence. 'Vacation' was the most positively valenced word in the data set, according to the ratings. 'Murders' was the seventh most negatively valenced word. The six words assigned more negatively valenced ratings were avoided on the basis that they were emotionally sensitive or offensive ('pedophile', 'rapist', 'AIDS', 'torture', 'leukemia', 'molester'). The simple trend lines shown in each graph all resembled the trend lines depicted in Fig. 2 and were based on the stimuli used by Pandey et al. [1]. The full range of graph stimuli shown to participants can be viewed at https://osf.io/5acjs/.

In sum, each participant saw the four trial types show in Fig. 2 (middle), presented in a randomized order. Each participant saw two graphs in which the situation in the displayed city was

Fig. 2. Left: Example of a trial completed by participants. All graphs (four per group) can be viewed at https://osf.io/5acjs. Middle: Trend (rising vs. falling) and valence (positive vs. negative) were varied within participants to create the four trial types shown to participants. Right: Each participant took part in one of six groups, each of which contained graphs that had been manipulated in a particular way. Axis Orientation (non-inverted vs. y-axis inverted vs. x-axis inverted) and Quantity Mapping (y-axis vs. x-axis) were varied between subjects to create these groups.

improving, and two graphs in which the situation was declining. Each of these graphs was labelled with a different fictional city name.

3.1.3 Participants

Access to Amazon Mechanical Turk requires users to be 18 years or older, restricting our sample to adult participants. Participation was restricted to US users of the platform, and payment was $1 per participant based on the US state of Illinois's minimum wage laws. Three hundred MTurk users volunteered to participate. Of this total, seven participants were excluded for answering an attention check question incorrectly prior to the study. Furthermore, three participants were excluded because their response time to one of the four trials was more than two standard deviations above the mean response time to all trials across all participants (>55.9 seconds). In total, data from 291 participants were included in the final analyses. For these participants, the age range was 24-72 years old (M = 39, SD = 11). In terms of gender, 130 participants were female (44.7%), 160 participants were male (55%), and one participant (0.3%) was non-binary/third gender.

3.1.4 Statistical analysis

All statistical analyses were conducted using statistical programming language R, version 4.0.3 [69], inside the integrated development environment RStudio 1.1.456 [70]. The packages 'plyr' 1.8.6 [71] and 'tidyverse' 1.3.0 [72] were used for data processing and visualization. The packages 'ggmcmc' 1.5.0 [73], and 'tidybayes' 2.3.1 [74] were used for data visualization. The package 'brms' 2.14.4 [75], [76] was used for Bayesian modelling.

The data and analysis scripts for this study are available at the publicly accessible Open Science Framework repository: https://osf.io/5acjs/. The hypotheses and planned analyses for this study were preregistered and can be found at the aforementioned URL. We used Bayesian multi-level logistic regression to assess our hypotheses. The dependent variable in all statistical models was Accuracy (incorrect vs correct). Accuracy corresponded to whether the participant correctly judged the situation in each city to be either improving or declining by selecting the appropriate multiple choice option. Valence was the only variable relevant to our hypotheses that varied within participants (each participant answered questions about graphs depicting both positively and negatively valenced quantities). To account for by-participant variation in Accuracy for Valence, random slopes and intercepts for Valence were included in all models. We set weakly informative priors centred at 0 with a standard deviation of 2.

The data from Experiment 1 were analysed using two statistical models. Model 1 assessed hypotheses 1–3 and included the variables Axis Orientation (inverted vs non-inverted), Quantity Mapping (x-axis vs y-axis), and Valence (negative vs positive). These variables were treatment-coded (0, 1). Model 2 assessed hypothesis 4 and included the variables Axis Orientation (inverted vs non-inverted) and Valence (negative vs positive), as well as the interaction between these variables. The variables in model 2 were deviation-coded (-0.5, +0.5) to facilitate interpretation of the main effects. Hypothesis 4 was assessed in a separate model because this hypothesis was relevant to only those graphs that plotted quantity along the y-axis.

For each of these models, we observed whether the 95% credible intervals for each variable included zero. Variables with a credible interval that did not contain zero were interpreted as strong evidence for the effect of this variable on accuracy.

3.1.5 Results

Table 2 shows the accuracy of responses to different graph stimuli in Experiment 1. Fig. 3 shows the coefficients for model 1 (top) and 2 (bottom). The blue posterior distributions show the range of log odds that represent credible values of the coefficient for each variable in the model. Values with a higher probability density, represented by denser areas of the distribution, are more credible than values with a lower probability density. The black lines beneath the distributions represent the 95% credible intervals for each variable.

Hypothesis 1 predicted that graphs without inverted axes would be easier to interpret than graphs with either the x- or y-axis inverted [1]. Consistent with this hypothesis, participants were more likely to respond correctly to graphs without inverted axes than to graphs with an inverted axis (see Table 2, rows 1-2). As shown in Fig. 3 (model 1), the credible interval for Axis Orientation did not include zero, indicating that there was strong evidence for the effect of this variable on the accuracy of responses.

Hypothesis 2 predicted that graphs plotting quantity information along the y-axis would be easier to interpret than graphs

TABLE 2
ACCURACY: EXPERIMENT 1

Graph Stimuli	Accuracy (Experiment 1)	
non-inverted	77.8%	*308 vs 87*
inverted	56.1%	*431 vs 337*
quantity on y-axis	55.0%	*317 vs 259*
quantity on x-axis	71.7%	*419 vs 165*
positive valence	72.9%	*423 vs 157*
negative valence	54.0%	*313 vs 267*
valence aligning: y-axis	66.7%	*192 vs 96*
valence misaligning: y-axis	43.4%	*125 vs 163*

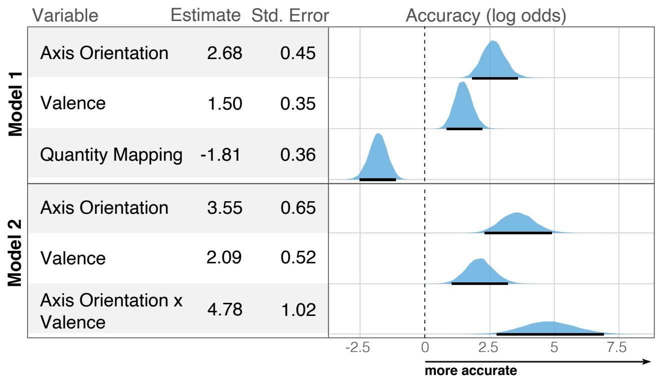


Fig. 3. Estimates, standard errors, and posterior distributions (blue) with 95% credible intervals (black lines) for coefficients in Bayesian logistic regression model 1 (all graphs) and 2 (graphs with quantity on y-axis) in Experiment 1.

plotting quantity information along the x-axis [44]. Contradicting this hypothesis, participants were more likely to respond correctly to graphs that represented quantity information along the x-axis rather than the y-axis (see Table 2, rows 3-4). As shown in Fig. 3 (model 1), the 95% credible interval for Quantity Mapping did not include zero, indicating that this effect was strongly supported by the data.

Hypothesis 3 predicted that graphs depicting a positively valenced quantity (vacation days) would be easier to interpret than graphs depicting a negatively valenced quantity (murders). Consistent with this hypothesis, participants were more likely to respond correctly to graphs depicting a positively valenced quantity than a negatively valenced quantity (see Table 2, rows 5-6). As shown in Fig. 3 (model 1), the 95% credible interval for Valence did not include zero, indicating that this effect was strongly supported by the data.

Hypothesis 4 predicted that graphs that aligned with vertical valence metaphors would be easier to interpret than graphs that did not align with these metaphors. Consistent with this hypothesis, participants were more likely to respond correctly to graphs that aligned with vertical valence metaphors compared to graphs that did not (see Table 2, rows 7-8). As shown in Fig. 3 (model 2), the 95% credible interval for Axis Orientation × Valence did not include zero, indicating that this effect was strongly supported by the data.

The main effects for model 2 predicted responses to be most accurate for non-inverted graphs depicting a positively valenced quantity (95% CI = [3.77, 7.08]). Non-inverted graphs depicting a negatively valenced quantity were predicted to elicit the second most accurate responses, although accuracy was substantially lower (95% CI = [-0.29, 2.08]). Predicted accuracy for inverted graphs depicting a positively valenced quantity (95% CI = [-1.39, 0.06]) and inverted graphs depicting a negatively valenced quantity (95% CI = [-1.16, 0.48]) were similar. These main effects suggest that Axis Orientation (inverted vs non-inverted) was the primary determinant of accuracy for graphs that mapped quantity information onto the y-axis, rather than Valence (negative vs positive) or Valence Alignment (misaligning vs aligning).

3.1.6 Discussion

Our results showed that participants were more likely to respond accurately to graphs without inverted axes compared to graphs with an inverted axis. This finding conceptually replicates the results reported by Pandey et al. [1] and extends these results to graphs with an inverted x-axis (rather than an inverted y-axis) and an inverted time axis (rather than an inverted quantity axis). Participants were also more likely to respond accurately to graphs that aligned with vertical valence metaphors compared to graphs that did not align with these metaphors. This result indicates that valence alignment matters for the interpretation of graphs that map quantity information onto the y-axis. Because our results show that inverting the axes of graphs generally has an adverse effect on response accuracy, a follow-up question is whether valence alignment matters enough to justify inverting the y-axis of graphs when these graphs depict negatively valenced quantities. Our results suggest not: responses to inverted graphs depicting negatively valenced quantities were less likely to be accurate than responses to non-inverted graphs depicting negatively valenced quantities. Ultimately, axis orientation with respect to the ordering of quantity along axes was a more important determinant of accuracy than valence alignment.

Our results also showed that participants were more likely to respond accurately to graphs depicting a positively valenced quantity than graphs depicting a negatively valenced quantity. This result is consistent with the cognitive processing advantage reported for positively valenced stimuli compared to negatively valenced stimuli [61], [62], [63]. In fact, for graphs that mapped quantity onto the y-axis, the experimental condition most likely to elicit accurate responses – by a substantial margin – was a non-inverted graph that depicted a positively valenced quantity (as shown by the main effects). The success of this experimental condition may be attributable to the additive effect of being non-inverted and depicting a positively valenced quantity, as well as the multiplicative effect of this experimental condition aligning with valence associations.

Participants were more likely to respond accurately to graphs mapping quantity information onto the x-axis compared to graphs mapping quantity information onto the y-axis. This result is surprising because line graphs conventionally represent quantity information using the y-axis, so the subversion of this convention may be expected to result in lower accuracy. However, the plotting of quantity onto the x-axis is a salient design choice that participants are likely to have spotted precisely due to its subversion of convention. The novelty of this design choice may have prompted participants to deliberate for longer when responding to this type of graph to verify that their answer was correct [77], [78]. In

TABLE 3
HYPOTHESIS 5 & 6: HORIZONTAL VALENCE ALIGNMENT

Stimuli	Valence Aligning (x-axis)	
	Left-handed	Right-handed
non-inverted axis, positive valence	no	yes
non-inverted axis, negative valence	yes	no
inverted axis, positive valence	yes	no
inverted axis, negative valence	no	yes

contrast, when the graph presented to participants mapped quantity onto its y-axis, consistent with graphical convention, participants may have been less likely to initiate a lengthy verification procedure, leaving room for errors to creep in. The latencies of participants' responses lend credence to this interpretation: participants spent less time responding to graphs that plotted quantity information using the y-axis (mean = 3.7 seconds) compared to graphs that plotted quantity information using the x-axis (mean = 5.5 seconds). However, it should be noted that there was little evidence for a general speed-accuracy trade-off in either experiment (see additional analyses at https://osf.io/5acjs/).

3.2 Experiment 2

Experiment 2 served as a replication-extension of Experiment 1. As a replication, we reproduced the experimental procedure from Experiment 1 and observed whether the new data exhibited the same patterns. As an extension, we introduced a test of horizontal valence metaphors. We investigated handedness-relative valence metaphors [47], which suggest that people associate the dominant side of their body (and therefore that side of space) with more positive valence, and vice versa. We also investigated the competing hypothesis that the representation of right-handed metaphors in cultural practices may lead even left-handers to associate positive valence with right-side space, a non-relative, 'absolute' metaphor [50], [51].

Experiment 2 tested the following two hypotheses in addition to Hypotheses 1–4 tested in Experiment 1:

H5. Graphs that align with handedness-relative horizontal valence metaphors will be easier to interpret than graphs that do not align with these metaphors (see Table 3).

H6. Graphs that align with absolute horizontal valence metaphors where 'good' is associated with rightward locations, irrespective of handedness, will be easier to interpret than graphs that do not align with these metaphors.

3.2.1 Methodology

The procedure for Experiment 2 followed a similar method to that detailed for Experiment 1. To assess our hypothesis regarding handedness-relative metaphors of emotional valence, eligibility criteria for participation specified that 50% of participants in each experimental condition (N = 25) should self-identify as left- and 50% as right-handed based on demographic data that participants provided to MTurk.

Three hundred MTurk users volunteered to participate. Of this total, six participants were excluded for answering an attention check question incorrectly prior to the study. Furthermore, seven participants were excluded because their response time to one of the trials was more than two standard deviations above the mean response time to all trials across all participants (25.9 seconds). Thus, data from 287 participants were included in the final analyses. The reported age range of participants was 23-72 years old (M = 39, SD = 10). In terms of gender, 175 participants were male (61%), 108 participants were female (37.6%), 2 participants were non-binary/third gender (0.7%), and 2 participants preferred not to say (0.7%). In terms of handedness, 145 participants were right-handed (50.5%) and 142 participants were left-handed (49.5%).

Three statistical models were used to analyse the data. Model 1 and 2 used in Experiment 1 were used again in Experiment 2. Model 3 assessed hypotheses 5–6 and contained the variables Axis Orientation (inverted vs non-inverted), Valence (negative vs positive), and Handedness (left vs right). This model also assessed the interaction between Axis Orientation and Valence (hypothesis 6), as well as the three-way interaction between Axis Orientation, Valence, and Handedness (hypothesis 5). The variables in model 3 were deviation-coded (-0.5, +0.5) to facilitate interpretation of the main effects. Hypotheses 5–6 were assessed in a separate model because these hypotheses applied only to graphs that mapped quantity information onto the x-axis.

3.2.2 Results

Table 4 shows the accuracy of responses to different graph stimuli in Experiment 1. Fig. 4 shows the coefficients for model 1 (top) and 2 (middle) and 3 (bottom). The blue posterior distributions represent the credible values of the coefficients for each variable and the probability density of these values. The black line beneath these distributions represents the 95% credible intervals.

Hypothesis 1 predicted that graphs without inverted axes would be easier to interpret than graphs with either the x- or y-axis inverted [1]. Consistent with this hypothesis, participants were more likely to respond correctly to graphs without inverted axes than to graphs with an inverted axis (see Table 4, rows 1-2). As shown in Fig. 4 (model 1), the 95% credible interval for Axis Orientation did not include zero, indicating that this effect was strongly supported by the data.

Hypothesis 2 predicted that graphs mapping quantity onto the y-axis would be easier to interpret than graphs mapping quantity onto the x-axis [44]. Contradicting this hypothesis, and in line with the results from Experiment 1, participants were more likely to respond correctly to graphs that represented quantity information along the x-axis than the y-axis (see Table 4, rows 3-4). As shown in Fig. 4 (model 1), the 95% credible interval for Quantity Mapping did not include zero, indicating that this effect was strongly supported by the data.

Hypothesis 3 predicted that graphs depicting a positively valenced quantity (vacation days) would be easier to interpret than graphs depicting a negatively valenced quantity (murders). There was a descriptive difference between graphs that depicted a positively valenced quantity and graphs that depicted a negatively valenced quantity (see Table 4, rows 5-6). However, as shown in Fig. 4 (model 1), the 95% credible interval for Valence marginally included zero, albeit with the bulk of the credible interval being above zero (lower bound = -0.10).

Hypothesis 4 predicted that graphs aligning with vertical valence metaphors would be easier to interpret than graphs misaligning with these metaphors. Consistent with this hypothesis, participants were more likely to respond correctly to graphs that aligned with vertical valence metaphors than to graphs that did not (see Table 4, rows 7-8). As shown in Fig. 4 (model 2), the 95%

TABLE 4
ACCURACY: EXPERIMENT 2

Graph Stimuli	Accuracy (Experiment 2)	
non-inverted	82.6%	*317 vs 67*
inverted	61.5%	*470 vs 294*
quantity on y-axis	59.7%	*346 vs 234*
quantity on x-axis	77.6%	*441 vs 127*
positive valence	71.6%	*411 vs 163*
negative valence	65.5%	*376 vs 198*
valence aligning: y-axis	64.5%	*187 vs 103*
valence misaligning: y-axis	54.8%	*159 vs 131*
valence aligning: x-axis (relative)	75.4%	*214 vs 70*
valence misaligning: x-axis (relative)	79.9%	*227 vs 57*
valence aligning: x-axis (absolute)	84.9%	*214 vs 43*
valence misaligning: x-axis (absolute)	70.4%	*200 vs 84*

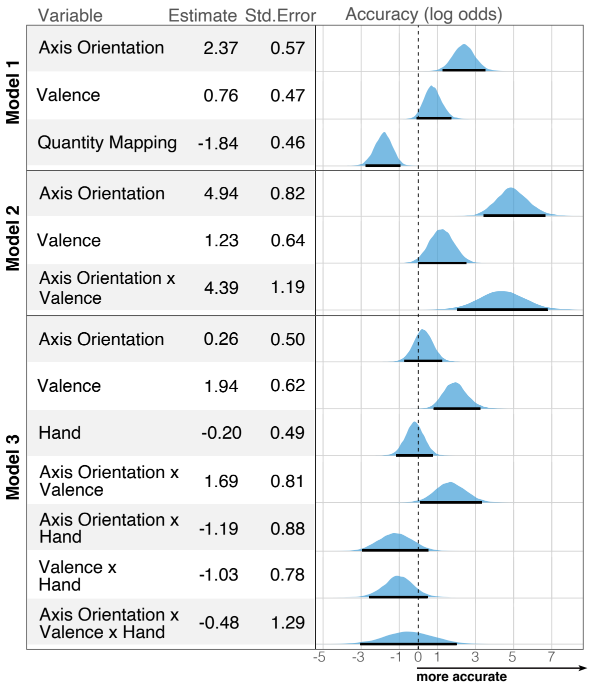


Fig. 4. Estimates, standard errors, and posterior distributions (blue) with 95% credible intervals (black lines) for coefficients in Bayesian logistic regression model 1 (all graphs), 2 (graphs with quantity on y-axis), and 3 (graphs with quantity on x-axis) in Experiment 2.

credible interval for Axis Orientation × Valence did not include zero, indicating that this effect was strongly supported by the data.

The main effects for model 2 predicted accuracy to be highest for non-inverted graphs depicting a positively valenced quantity (95% *CI* = [4.53, 8.44]), followed by non-inverted graphs depicting a negatively valenced quantity (95% *CI* = [1.47, 4.71]). Accuracy for inverted graphs depicting a negatively valenced quantity (95% *CI* = [-0.83, 1.27]) and inverted graphs depicting a positively valenced quantity (95% *CI* = [-1.73, 0.22]) were predicted to be substantially lower. These main effects indicate that Axis Orientation (inverted vs normal) was the primary determinant of accuracy for graphs that plotted quantity using the y-axis, rather than Valence (negative vs positive) or Valence Alignment (misaligning vs aligning).

Hypothesis 5 predicted that graphs that align with handedness-relative horizontal valence metaphors would be easier to interpret than graphs that do not align with these metaphors. There was a descriptive difference in accuracy between graphs that did or did not align with handedness-relative horizontal valence metaphors (see Table 4, rows 9-10). However, as shown in Fig. 4 (model 3), the 95% credible interval for Axis Orientation × Valence × Hand included zero, indicating that this effect was not strongly supported by the data.

Hypothesis 6 predicted that graphs that align with absolute horizontal valence metaphors, irrespective of participants' handedness, would be easier to interpret than graphs that do not align with these metaphors. Consistent with this hypothesis, graphs that aligned with the absolute vertical valence metaphor were more likely to elicit correct responses than graphs that did not align with this metaphor (see Table 4, rows 11-12). As shown in Fig. 4 (model 3), the 95% credible interval for Axis Orientation × Valence did not include zero, indicating that this effect was strongly supported by the data.

The main effects for model 3 predicted accuracy to be highest for non-inverted graphs depicting a positively valenced quantity (95% *CI* = [2.64, 5.68]). The second highest predicted accuracy was elicited by inverted graphs depicting a positively valenced quantity (95% *CI* = [1.85, 4.23]). Accuracy for inverted graphs depicting a negatively valenced quantity (95% *CI* = [1.11, 2.61]) and non-inverted graphs depicting a negatively valenced quantity (95% *CI* = [0.39, 2.15]) were similar. These main effects indicate that Valence (negative vs positive) was the primary determinant of accuracy for graphs that plotted quantity using the x-axis, rather than Axis Orientation (inverted vs non-inverted) or Valence Alignment (misaligning vs aligning).

3.2.3 *Discussion*

Consistent with Experiment 1 and with the results of Pandey et al. [1], our results show that participants were more likely to respond accurately to graphs without inverted axes than to graphs with inverted axes. Participants were also more likely to respond accurately to graphs that aligned with vertical valence metaphors than to graphs that did not align with these metaphors. For graphs that depicted quantity information using the y-axis, axis orientation mattered more for graph interpretability than alignment with vertical valence metaphors. For instance, accuracy was higher for non-inverted graphs depicting a negatively valenced quantity than for inverted graphs depicting a negatively valenced quantity, despite the fact that the latter graph aligned with vertical valence metaphors. Altogether, these results indicate that alignment with valence metaphors matters for the interpretation of line graphs, but not enough to justify inverting the axis of these graphs, which is likely to confuse graph users [1], [35].

Unlike Experiment 1, the credible interval for emotional valence (negative vs positive) in Experiment 2 included zero. However, the majority of this credible interval was above zero, with the lower bound of the interval being -0.10. Furthermore, while it was not the primary test of our hypothesis, the model assessing only horizontal valence metaphors indicated that there was strong evidence for an effect of emotional valence for graphs that depicted quantity using the x-axis (95% *CI* = [0.81, 3.25]). Therefore, data across both experiments could be interpreted together as showing an effect of emotional valence on response accuracy, consistent with research showing that positively valenced stimuli are recognised more quickly than negatively valenced stimuli [61], [62], [63].

As in Experiment 1, participants in Experiment 2 were more likely to respond accurately to graphs that mapped quantity information onto the x-axis rather than the y-axis (in contrast to [44]). Participants also spent more time responding to these graphs (mean = 5.11 seconds) than to graphs that depicted quantity information using the y-axis (mean = 3.79 seconds). This result lends further support to the notion that participants made fewer mistakes when responding to x-axis graphs because they spent longer verifying that their answers to these graphs were correct [77], [78].

For graphs mapping quantity onto the x-axis, participants were more likely to respond accurately to graphs that aligned with absolute valence metaphors (where 'good' is associated with rightward space) rather than handedness-relative valence metaphors (where 'good' is associated with the side of space corresponding to the dominant side of one's body). This finding supports the idea that the representation of the right-handed valence metaphor in cultural practices and language causes an absolute association between right-side space and positive valence that extends even to left-handers [50], [51].

Our results also showed that emotional valence (negative vs positive) mattered more than axis orientation for the interpretability of graphs mapping quantity onto the x-axis. In fact, accuracy was similar for x-axis graphs with non-inverted axes (78.2%, 147 vs 41) compared to x-axis graphs with an inverted axis (77.4%, 294 vs 86). A possible reason for the lack of an axis orientation effect for x-axis graphs may be that the novel presentation of quantity on the x-axis caused participants to pay more attention to the direction of this axis. The lack of an axis orientation effect may then have left more room for emotional valence to influence responses.

4 General Discussion

Our main results can be summarised as follows:

- Non-inverted line graphs are easier to interpret than line graphs with an inverted axis, confirming and extending results reported by Pandey et al. [1] to graphs with an inverted x-axis, and graphs with an inverted time axis (either the y-axis or x-axis).
- Line graphs that adhere to conceptual metaphors of emotional valence [19], [20], [21] are easier to interpret than line graphs that do not adhere to these conceptual metaphors. However, the beneficial effect of valence alignment on graph interpretability does not outweigh the adverse effect of reversing graphical axes.
- Line graphs depicting a positively valenced quantity are easier to interpret than line graphs depicting a negatively valenced quantity, both when quantity was mapped onto the y-axis and the x-axis, consistent with the negativity bias [61], [63], [65].
- Line graphs that plot quantity onto the x-axis are easier to interpret than line graphs that plot quantity onto the y-axis.

Our results support the idea that conceptual metaphor theory can be relevant to the design of data visualizations [7], [8] – in this case, line graphs. However, our results lead us to recommend that line graphs should be designed to align with conceptual metaphors of emotional valence only if doing so does not subvert other well-established design graphical conventions. Our results also suggest that the inverted y-axis of the Stand Your Ground visualization, which depicted gun deaths in Florida [26], did indeed have the potential to mislead viewers of the visualization. Contrary to comments made by graphic designer Christine Chan, while axis direction *can* "be shown either way" [26], we recommend that designers hoping to elicit the correct interpretation of the data being visualized should not invert the y-axis.

There are, however, some visualizations in which y-axis inversion seems to have been employed successfully [36]. For example, the infographic entitled Iraq's Bloody Toll (see Fig. 1, B), designed by Simon Scarr, was awarded the silver prize at annual infographics conference Malofiej [17], [79]. This infographic depicts the number of deaths per month in Iraq between 2003 and 2011, inverting its y-axis and filling its bars in red to resemble blood running down the page. In contrast to the Stand Your Ground visualization, the Iraq infographic's title and its use of thin red bars with rounded end caps emphasise the metaphor of blood. In addition, the use of bars in the Iraq infographic allows the viewer to see that the bars originate at the top of the page, whereas the Stand Your Ground graph is a line chart and so does not provide an easy cue for determining axis orientation [36]. More research is needed, but y-axis inversion may be acceptable if the axis inversion is made highly salient by other design features of the visualization. However, doing so necessarily involves risk as some viewers might not notice the axis inversion.

In some cases, graph designers may prefer to avoid plotting quantity on the y-axis entirely if using the y-axis for quantity conflicts with metaphors of emotional valence, and instead may prefer to use the x-axis. In support of this recommendation, Tversky [2] has argued that the horizontal axis is more neutral than the vertical one. Our results corroborate the relative neutrality of the horizontal axis with respect to line graphs: while we found an effect of valence alignment for the x-axis in Experiment 2, this effect was not as strong as the effect of valence alignment for the vertical axis in both Experiment 1 and 2 (as indicated by the distance of the 95% credible intervals from zero, see Fig. 3, and 4).

Our results have potential applications to other graphs where designers select the ordering of categorical variables, such as stacked bar charts. When there is no strict convention for the ordering of y-axis categories, these categories could be listed from smallest to largest from the bottom to the top of the y-axis, or from the top to the bottom, presumably without causing confusion. The interpretability of these graphs might be enhanced when the category with the 'most' is positioned at the top of the graph, and the category with the 'least' is positioned at the bottom, in line with vertical quantity metaphors [3], [4], [6]. Alternatively, the optimal ordering might depend on whether the quantity is positively or negatively valenced [20], [21], [38]. This is an intriguing avenue of research for future studies to explore.

In general, more research needs to be conducted looking at graphs showing different trends and quantities to determine the extent to which our results can be generalised. For example, the graphs we showed to participants were based on those used by Pandey et al. [1] and thus were highly schematic, showing a simple, relatively linear line without any other fluctuations. Real-life data are often more noisy and complex than in the current work. Future research is needed to determine whether our results generalize to different types of data, specifically data with more variability. It also remains to be seen how non-linearities in the trend line might affect conceptualisations of emotional valence.

In addition to the vertical axis, our data revealed an effect of absolute horizontal metaphors of emotional valence that was not relative to handedness. Interpretation of the line graphs in this study thus did not appear to be influenced by body-specificity [47], [48], [49]. This finding supports the notion that even left-handers may come to acquire an absolute association between right-side space and positive valence due to the representation of the right-handed valence metaphor in cultural practices and language [50], [51].

Based on the support our results provided for the negativity bias, an additional recommendation we can make for the design of line graphs is that, if it is possible to frame a quantity in positively or negatively valenced terms (e.g., '10% less sad' vs '10% happier', or '30% uncertain' vs '70% certain'), the visualization designer should use the positively valenced framing. For example, it may be easier to interpret the graph shown in Fig. 1 (C) if this graph displayed ratings of favourability rather than unfavourability.

A possible criticism of our assessment of emotional valence is that participants may not have universally agreed that 'vacation days' is a positively valenced concept – the judgment of a stimulus as positively or negatively valenced is subjective. For example, employers might view 'vacation days' as negatively valenced due to its association with diminished workplace productivity. Nonetheless, the fact that 'vacation' was rated as the most positively valenced word in the 13,915 words normed by Warriner et al. [68] leads us to assume that the overwhelming majority of our participants assessed this concept favourably. Because Warriner et al. used a sample of US participants recruited via Amazon Mechanical Turk, the majority of whom had English as their native language, similar to our sample, there is good reason to believe that our participants shared similar views about the valence of 'vacation days'. Furthermore, the standard deviation of the ratings for 'vacation' (0.77) was substantially below the average standard deviation for words in the Warriner et al. dataset as a whole (1.68) (more extreme ratings tend to be less variable, see [80]), indicating that attitudes toward 'vacation' were highly consistent. The large samples in both our experiments ensured that anomalous data from participants who felt differently did not inordinately skew our results. Finally, because participants answered questions about both vacation days and murders, it seems likely that vacation days would have been perceived as more positively valenced than murders in relative terms. This contrast in itself might be expected to affect results in the direction we observed.

Our finding that participants were more likely to respond accurately to graphs that plotted quantity using the x-axis was contrary to our predictions. We speculated that the subversion of convention in these graphs was likely to have been salient to participants, who then spent longer interpreting these graphs to make sure their response was correct. This finding differs from the y-axis advantage for bar charts reported by Fischer et al. [44]. The discrepancy between these two sets of results may be attributable to the fact that horizontal bar charts are relatively 'normal', conventional visualizations, unlike line graphs that map quantity onto the x-axis, which are unusual. Our results may also speak to the notion of *desirable difficulties*, which posits that making learning more challenging can be beneficial [77], [78].

This paper demonstrates the role that the field of conceptual metaphor research may have to play in the design of data visualizations. Our results suggest that conceptual metaphor theory matters for the interpretation of line graphs. However, visualization designers are warned against subverting graphical convention in order to align with conceptual metaphors that are not conventionally instantiated graphically. Otherwise, the resultant visualization may be more liable to mislead than to inform.

Acknowledgments

Greg Woodin was supported by the Economic and Social Research Council [grant number ES/P000711/1]. Bodo Winter was supported by a UKRI Future Leaders Fellowship MR/T040505/1.

References

[1] A. V. Pandey, K. Rall, M. L. Satterthwaite, O. Nov, and E. Bertini, "How deceptive are deceptive visualizations?: An empirical analysis of common distortion techniques," *New York University Public Law and Legal Theory Working Papers*, no. 504, 2015.

[2] B. Tversky, "Visualizing thought," *Topics in Cognitive Science*, vol. 3, no. 3, pp. 499–535, 2011.

[3] M. Hartmann, V. Gashaj, A. Stahnke, and F. W. Mast, "There is more than "more is up": Hand and foot responses reverse the vertical association of number magnitudes," *Journal of Experimental Psychology. Human Perception and Performance*, vol. 40, no. 4, pp. 1401–1414, 2014.

[4] B. Winter and T. Matlock, "More is up... and right: Random number generation along two axes," in *Proceedings of the Annual Meeting of the Cognitive Science Society*. Austin, Texas: Cognitive Science Society, 2013, pp. 3789–3974.

[5] T. Loetscher, C. J. Bockisch, M. E. R. Nicholls, and P. Brugger, "Eye position predicts what number you have in mind," *Current Biology*, vol. 20, no. 6, pp. R264–265, 2010.

[6] M. Hartmann, L. Grabherr, and F. W. Mast, "Moving along the mental number line: Interactions between whole-body motion and numerical cognition," *Journal of Experimental Psychology. Human Perception and Performance*, vol. 38, no. 6, pp. 1416–1427, 2012.

[7] P. Parsons, "Conceptual Metaphor Theory as a foundation for communicative visualization design," in *Visualization for Communication (VisComm)*. Berlin, Germany: IEEE VIS, 2018.

[8] J. S. Risch, "On the role of metaphor in information visualization," *arXiv*, 2008. [Online]. Available: http://arxiv.org/abs/0809.0884

[9] R. W. Gibbs, Jr., *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge, UK: Cambridge University Press, 1994.

[10] Z. Kövecses, *Metaphor: A Practical Introduction*. New York City, New York: Oxford University Press, 2002.

[11] G. Lakoff and M. Johnson, *Metaphors We Live By*. Chicago, Illinois: University of Chicago Press, 1980.

[12] T. Loetscher, C. J. Bockisch, and P. Brugger, "Looking for the answer: The mind's eye in number space," *Neuroscience*, vol. 151, no. 3, pp. 725–729, 2008.

[13] L. Boroditsky, "Metaphoric structuring: understanding time through spatial metaphors," *Cognition*, vol. 75, no. 1, pp. 1–28, 2000.

[14] L. Boroditsky and M. Ramscar, "The roles of body and mind in abstract thought," *Psychological Science*, vol. 13, no. 2, 2002.

[15] M. S. McGlone and J. L. Harding, "Back (or forward?) to the future: The role of perspective in temporal language comprehension," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 24, no. 5, pp. 1211–1223, 1998.

[16] Reuters, 2014. [Online]. Available: https://www.reuters.com/news/world

[17] S. Scarr, "Iraq's bloody toll," 2011. [Online]. Available: https://www.scmp.com/infographics/article/1284683/iraqs-bloody-toll

[18] H. Enten, "Americans' distaste for both Trump and Clinton is record-breaking," 2016. [Online]. Available: https://fivethirtyeight.com/features/americans-distaste-for-both-trump-and-clinton-is-record-breaking/

[19] B. P. Meier and M. D. Robinson, "Does "feeling down" mean seeing down? Depressive symptoms and vertical selective attention," *Journal of Research in Personality*, vol. 40, no. 4, 2006.

[20] ——, "Why the sunny side is up: Association between affect and vertical position," *Psychological Science*, vol. 15, no. 4, pp. 243–247, 2004.

[21] D. Casasanto and K. Dijkstra, "Motor action and emotional memory," *Cognition*, vol. 115, no. 1, pp. 179–185, 2010.

[22] L. F. Barrett and J. A. Russell, "The structure of current affect: Controversies and emerging consensus," *Current Directions in Psychological Science*, vol. 8, no. 1, 2016.

[23] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "Motivated attention: affect, activation and action," in *Attention and Orienting: Sensory and Motivational Processes*, P. J. Lang, R. F. Simons, and M. T. Balaban, Eds. Lawrence Erlbaum Associates Publishers, 1997.

[24] J. A. Russell, "Core affect and the psychological construction of emotion," *Psychological Review*, vol. 110, no. 1, pp. 145–172, 2003.

[25] P. Engel, *Gun deaths in Florida increased with 'Stand Your Ground'*, 2014. [Online]. Available: https://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2?r=US&IR=T

[26] M. Lallanilla, *Misleading gun-death chart draws fire*, 2014. [Online]. Available: https://www.livescience.com/45083-misleading-gun-death-chart.html

[27] J. deVilla, *Lies, damned lies, and Reuters' graph of Florida firearm deaths before and after "Stand Your Ground"*, 2014. [Online]. Available: http://www.joeydevilla.com/2014/04/14/lies-damned-lies-and-reuters-graph-of-florida-firearm-deaths-before-and-after-stand-your-ground

[28] L. Wade, *How to lie with statistics: Stand Your Ground and gun deaths*, 2014. [Online]. Available: https://thesocietypages.org/socimages/2014/12/28/how-to-lie-with-statistics-stand-your-ground-and-gun-deaths/

[29] L. M. Padilla, "A case for cognitive models in visualization research," in *2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization (BELIV)*. Berlin, Germany: IEEE, 2018, pp. 69–77.

[30] L. M. Padilla, S. H. Creem-Regehr, M. Hegarty, and J. K. Stefanucci, "Decision making with visualizations: A cognitive framework across disciplines," *Cognitive Research: Principles and Implications*, vol. 3, no. 29, pp. 1–25, 2018.

[31] M. Ishihara, P. E. Keller, Y. Rossetti, and W. Prinz, "Horizontal spatial representations of time: evidence for the STEARC effect," *Cortex*, vol. 44, no. 4, pp. 454–61, 2008.

[32] D. J. Cox, "The art and science of visualization: metaphorical maps and cultural models," *Technoetic Arts*, vol. 2, no. 2, pp. 71–80, 2004.

[33] P. A. Fishwick, "Metaphoric mappings: The art of visualization," in *Aesthetic Computing*, P. A. Fishwick, Ed. MITP, 2006, pp. 89–114.

[34] I. Vessey, "Cognitive fit: A theory-based analysis of the graphs versus tables literature*," *Decision Sciences*, vol. 22, no. 2, pp. 219–240, 1991.

[35] M. Correll and J. Heer, "Black Hat Visualization," in *DECISIVe: Workshop on Dealing with Cognitive Biases in Visualisations*. IEEE, 2017.

[36] R. Kosara, "When bars point down," 2014. [Online]. Available: https://eagereyes.org/journalism/when-bars-point-down

[37] M. Hartmann, F. W. Mast, and M. H. Fischer, "Counting is a spatial process: Evidence from eye movements," *Psychological Research*, vol. 80, no. 3, pp. 399–409, 2016.

[38] T. Seno, T. Kawabe, H. Ito, and S. Sunaga, "Vection modulates emotional valence of autobiographical episodic memories," *Cognition*, vol. 126, no. 1, pp. 115–120, 2013.

[39] B. Winter and T. Matlock, "Primary metaphors are both cultural and embodied," in *Metaphor: Embodied Cognition and Discourse*, B. Hampe, Ed. Cambridge, UK: Cambridge University Press, 2017, pp. 99–116.

[40] A. Chinello, M. D. de Hevia, C. Geraci, and L. Girelli, "Finding the spatial-numerical association of response codes (SNARC) in signed numbers: Notational effects in accessing number representation," *Functional Neurology*, vol. 27, no. 3, pp. 177–185, 2013.

[41] S. Dehaene, S. Bossini, and P. Giraux, "The mental representation of parity and number magnitude," *Journal of Experimental Psychology: General*, vol. 122, no. 3, p. 371, 1993.

[42] G. Wood, K. Willmes, H.-C. Nuerk, and M. H. Fischer, "On the cognitive link between space and number: A meta-analysis of the SNARC effect," *Psychology Science*, vol. 50, no. 4, pp. 489–525, 2008.

[43] G. Woodin and B. Winter, "Placing abstract concepts in space: Quantity, time and emotional valence," *Frontiers in Psychology*, vol. 9, pp. 1–14, 2018.

[44] M. H. Fischer, N. Dewulf, and R. L. Hill, "Designing bar graphs: orientation matters," *Applied Cognitive Psychology*, vol. 19, no. 7, pp. 953–962, 2005.

[45] B. Winter, T. Matlock, S. Shaki, and M. H. Fischer, "Mental number space in three dimensions," *Neuroscience and Biobehavioral Reviews*, vol. 57, pp. 209–219, 2015.

[46] B. Winter, T. Marghetis, and T. Matlock, "Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number," *Cortex*, vol. 64, pp. 209–224, 2015.

[47] D. Casasanto, "Embodiment of abstract concepts: Good and bad in right- and left-handers," *Journal of Experimental Psychology: General*, vol. 138, no. 3, pp. 351–367, 2009.

[48] D. Casasanto and T. Henetz, "Handedness shapes children's abstract concepts," *Cognitive Science*, vol. 36, no. 2, pp. 359–372, 2012.

[49] J. W. Van Strien and S. Van Beek, "Ratings of emotion in laterally presented faces: Sex and handedness effects," *Brain and Cognition*, vol. 44, no. 3, pp. 645–652, 2000.

[50] D. Casasanto, "The hierarchical structure of mental metaphors," in *Metaphor: Embodied Cognition and Discourse*, B. Hampe, Ed. Cambridge, UK: Cambridge University Press, 2017, pp. 46–61.

[51] D. Casasanto and R. Bottini, "Spatial language and abstract concepts," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 5, no. 2, pp. 139–149, 2014.

[52] S. Y. Kim, J. W. Lee, and I. S. Sohn, "Comparison of various statistical methods for identifying differential gene expression in replicated microarray data," *Statistical Methods in Medical Research*, vol. 15, no. 1, pp. 3–20, 2016.

[53] A. Abdul-Rahman, S. Walton, K. Bemis, J. G. Lein, K. Coles, D. Silver, and M. Chen, "Re-spatialization of time series plots," *Information Visualization*, vol. 16, no. 4, pp. 275–290, 2017.

[54] S. Havre, E. Hetzler, P. Whitney, and L. Nowell, "ThemeRiver: Visualizing thematic changes in large document collections," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 1, pp. 9–20, 2002.

[55] M. Wattenberg, "Baby names, visualization, and social data analysis," in *IEEE Symposium on Information Visualization*. Minneapolis, Minnesota: IEEE, 2005, pp. 1–7.

[56] A. Vallesi, M. A. Binns, and T. Shallice, "An effect of spatial-temporal association of response codes: Understanding the cognitive representations of time," *Cognition*, vol. 107, no. 2, pp. 501–527, 2008.

[57] M. Conson, F. Cinque, A. M. Barbarulo, and L. Trojano, "A common processing system for duration, order and spatial information: Evidence from a time estimation task," *Experimental Brain Research*, vol. 187, no. 2, pp. 267–274, 2008.

[58] S. Ruiz Fernandéz, M. Lachmair, and J. J. Rahona, "Human mental representation of time in the vertical space," in *Proceedings of the 6th International Congress of Medicine in Space and Extreme Environments*. Berlin, Germany: 6th International Congress of Medicine in Space and Extreme Environments, 2014.

[59] M. J. Leone, A. Salles, A. Pulver, D. A. Golombek, and M. Sigman, "Time drawings: Spatial representation of temporal concepts," *Consciousness and Cognition*, vol. 59, pp. 10–25, 2018.

[60] M. H. Fischer and H. Campens, "Pointing to numbers and grasping magnitudes," *Experimental Brain Research*, vol. 192, no. 1, pp. 149–153, 2009.

[61] M. W. Matlin and D. J. Stang, *The Pollyanna Principle: Selectivity in Language, Memory, and Thought*. Cambridge, Massachusetts: Schenkman Pub. Co., 1978.

[62] G. Peeters, "Evaluative positive-negative asymmetry in adjective-noun compounds," *Polish Psychological Bulletin*, vol. 20, no. 4, pp. 255–266, 1989.

[63] ——, "The positive-negative asymmetry: On cognitive consistency and positivity bias." *European Journal of Social Psychology*, vol. 1, pp. 455–474, 1971.

[64] G. Peeters and J. Czapinski, "Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects," *European Review of Social Psychology*, vol. 1, no. 1, pp. 33–60, 1990.

[65] A. Vaish, T. Grossmann, and A. Woodward, "Not all emotions are created equal: The negativity bias in social-emotional development," *Psychological Bulletin*, vol. 134, no. 3, pp. 383–403, 2008.

[66] Qualtrics, *Four core experiences of business: Four applications to get them right*, 2019. [Online]. Available: https://www.qualtrics.com/

[67] Amazon Mechanical Turk, 2021. [Online]. Available: https://www.mturk.com/

[68] A. B. Warriner, V. Kuperman, and M. Brysbaert, "Norms of valence, arousal, and dominance for 13,915 English lemmas," *Behavior Research Methods*, vol. 45, no. 4, pp. 1191–1207, 2013.

[69] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2020. [Online]. Available: https://www.R-project.org/

[70] RStudio Team, *RStudio: Integrated Development Environment for R*, RStudio, Inc., Boston, MA, 2016. [Online]. Available: http://www.rstudio.com/

[71] H. Wickham, "The split-apply-combine strategy for data analysis," *Journal of Statistical Software*, vol. 40, no. 1, pp. 1–29, 2011.

[72] H. Wickham, M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and

H. Yutani, “Welcome to the tidyverse,” *Journal of Open Source Software*, vol. 4, no. 43, p. 1686, 2019.

[73] X. F. i Marín, “ggmcmc: Analysis of MCMC samples and Bayesian inference,” *Journal of Statistical Software*, vol. 70, no. 9, pp. 1–20, 2016.

[74] M. Kay, *tidybayes: Tidy Data and Geoms for Bayesian Models*, 2020, r package version 2.3.1. [Online]. Available: http://mjskay.github.io/tidybayes/

[75] P.-C. Bürkner, “brms: An R package for Bayesian multilevel models using Stan,” *Journal of Statistical Software*, vol. 80, no. 1, pp. 1–28, 2017.

[76] ——, “Advanced Bayesian multilevel modeling with the R package brms,” *The R Journal*, vol. 10, no. 1, pp. 395–411, 2018.

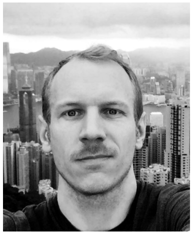
[77] R. A. Bjork, “Memory and metamemory considerations in the training of human beings,” in *Metacognition: Knowing about Knowing*, A. P. Metcalfe, J. Shimamura, Ed. Cambridge, Massachusetts: The MIT Press, 1994, pp. 185–205.

[78] J. Hullman, E. Adar, and P. Shah, “Benefitting InfoVis with visual difficulties,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2213–2222, 2011.

[79] A. Cotgreave, “Lies, damn lies, and statistics: How to take something positive from the UK’s EU referendum campaign,” 2016. [Online]. Available: https://www.infoworld.com/article/3088166/why-how-to-lie-with-statistics-did-us-a-disservice.html

[80] L. Pollock, “Statistical and methodological problems with concreteness and other semantic variables: A list memory experiment case study,” *Behavior Research Methods*, vol. 50, no. 3, pp. 1198–1216, 2018.

Greg Woodin is a PhD student in English Language and Linguistics at the University of Birmingham, UK. His research interests include metaphor, iconicity, gesture, spatial cognition, and data visualization.



Bodo Winter is a Senior Lecturer in cognitive linguistics at the University of Birmingham, UK, and an UKRI Future Leaders Fellow. His research interests include metaphor, gesture, language evolution, and statistical methodology.

Lace Padilla is an Assistant Professor in the Department of Cognitive & Information Sciences at UC Merced, USA. Her research interests include visual decision-making and spatial cognition.