PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Winter B. 2022 Abstract concepts and emotion: cross-linguistic evidence and arguments against affective embodiment. *Phil. Trans. R. Soc. B* **378**: 20210368. https://doi.org/10.1098/rstb.2021.0368

Received: 29 October 2021
Accepted: 7 February 2022

One contribution of 23 to a theme issue 'Concepts in interaction: social engagement and inner experiences'.

Subject Areas:
neuroscience, cognition

Keywords:
abstract concepts, embodied cognition, semantic memory, emotion, concreteness, embodiment

Author for correspondence:
Bodo Winter
e-mail: b.winter@bham.ac.uk

THE ROYAL SOCIETY PUBLISHING

Abstract concepts and emotion: cross-linguistic evidence and arguments against affective embodiment

Bodo Winter

Department of English Language and Linguistics, University of Birmingham, Birmingham B15 2TT, UK

BW, 0000-0001-6036-6774

How are abstract concepts such as 'freedom' and 'democracy' represented in the mind? One prominent proposal suggests that abstract concepts are grounded in emotion. Supporting this 'affective embodiment' account, abstract concepts are rated to be more strongly positive or more strongly negative than concrete concepts. This paper demonstrates that this finding generalizes across languages by synthesizing rating data from Cantonese, Mandarin Chinese, Croatian, Dutch, French, German, Indonesian, Italian, Polish and Spanish. However, a deeper look at the same data suggests that the idea of emotional grounding only characterizes a small subset of abstract concepts. Moreover, when the concreteness/abstractness dimension is not operationalized using concreteness ratings, it is actually found that concrete concepts are rated as more emotional than abstract ones. Altogether, these results suggest limitations to the idea that emotion is an important factor in the grounding of abstract concepts.

This article is part of the theme issue 'Concepts in interaction: social engagement and inner experiences'.

1. Introduction

One of the most important distinctions in the study of concepts is that between abstract and concrete concepts. Words for abstract concepts are generally processed more slowly [1–10], memorized less accurately [11–14] and acquired later than words for concrete concepts [15–18]. Although there is controversy about what exactly characterizes the concreteness/abstractness divide [19–21], most researchers consider concepts to be abstract 'if they do not apply to physical objects that we can touch, see, feel, hear, smell or taste' [19, p. 1]. Owing to their lack of sensorimotor content, such concepts are generally seen as a challenge to grounded or embodied theories of cognition [22,23]. These theories posit that comprehending a concept involves performing mental simulations that engage sensorimotor systems in the brain [24–28], such as activating leg-related neurons when processing the verbal concept 'kick' [29,30]. But, at least at first sight, sensorimotor simulation seems to be unable to explain how abstract concepts are comprehended [23,31–33]. After all, how could abstract concepts such as 'virtue', 'idea' or 'democracy' involve the activation of sensorimotor systems when the content of these concepts has so little to do with anything that can be directly perceived or acted upon in a physical manner?

Accepting 'the challenge of abstract concepts' [22], researchers within the embodied tradition have proposed what some call 'multiple representation' theories of abstract concepts [34,35], perhaps most prominently expressed in the 'Words as social tools' (WAT) account by Borghi *et al.* [36]. Based on this view, abstract concepts are supported by multiple distinct cognitive systems. The most dominant cognitive system supporting abstract concepts is generally thought to be language, with virtually any proposal about the nature of abstract concepts acknowledging the importance of linguistic or language-like representations [36–42]. Decades-worth of research under the banner of Paivio's dual coding theory suggests that abstract concepts rely more exclusively on a verbal mode of representation than do concrete concepts [43–47]. Neuro-imaging research finds enhanced haemodynamic activity for abstract as opposed to concrete concepts in left-hemisphere language networks [1,12,48–50]. In addition, language competence predicts improved processing of abstract words in children [51], and the richness of linguistic contexts facilitates abstract concepts more than concrete ones [52]. Thus, a large body of behavioural and neuroscientific evidence points to language as a key factor in the representation and processing of abstract concepts.

Embodied cognition researchers, however, generally follow the supposition that all concepts, including abstract ones, have to have some level of 'grounding' in systems external to language. This has led to proposals that abstract concepts are supported by interoception [53], simulations of situated experiences [21,54], social metacognition [36,55] and conceptual metaphor [32]. Another influential proposal argues that abstract concepts make up for their lack of sensorimotor content by virtue of being represented in terms of affective or emotional content [16,17,39,51,56–59]. Yao *et al.* [7] describe this account as one of 'representational substitution', a term that encapsulates the idea that affective representations may *substitute* for the sensorimotor representations that abstract concepts lack. This idea has also been dubbed 'emotional grounding' [4], 'affective grounding' [60] or 'affective embodiment' [36]. Henceforth, I shall use 'emotional grounding' as a cover term.

This paper will present evidence that at first sight seems to support the emotional grounding of abstract concepts by generalizing the account to data from new languages, adding a much-needed cross-linguistic dimension to a literature that is largely focused on abstract concepts in English. However, a closer look at the same data reveals that the idea of emotional grounding characterizes, if at all, only a very small minority of concepts. Much of the recent discussion on abstract concepts has argued that we need to recognize that abstract concepts as a group are characterized by heterogeneity [19,55,61,62]. For example, a recent cluster analysis of semantic rating data found evidence for at least four kinds of abstract concepts: (i) philosophical and spiritual, (ii) self and sociality, (iii) emotion and inner states, and (iv) physical, spatio-temporal and quantitative concepts [55]. As stated by Pecher [31, p. 502], 'abstract concepts should not be considered as a single kind of concept'. Fully in line with this emerging recognition that there are distinct varieties of abstract concepts, this paper argues that emotional grounding has been erroneously predicated on *all* abstract concepts. Finally, in line with many critiques of concreteness ratings [19,20,63–65], this paper empirically demonstrates that once we operationalize concreteness differently, the relationship between emotional valence and concreteness reverses sign, i.e. we obtain results that are diametrically opposed to the emotional grounding hypothesis.

2. Background. How does emotion relate to the concrete/abstract divide?

As has been discussed in existing reviews of the topic, the evidence for emotional grounding is mixed ([22, pp. 273–274]; see also [7]). This section reviews evidence for and against emotional grounding. I shall discuss studies that look at the effects of emotional valence differentially in concrete and abstract concepts, regardless of whether these are studies focused on conceptual structure, processing or acquisition. Even though the present study does not itself investigate language acquisition, results on the acquisition of words/concepts are relevant because much of the recent literature on emotional grounding has focused on the idea that emotions may be part of a 'bootstrapping' mechanism for the acquisition of abstract concepts. As expressed by Ponari *et al.* [17, p.2], 'while words referring to concrete objects and actions would be learnt by associating sensory-motor experience with the word, abstract words would be learned by associating emotional states with the word'. Thus, acquisition results need to be reviewed to get a full picture of the evidence for and against emotional grounding. Most generally, results where emotional valence plays a stronger role in representation, processing or acquisition specifically for abstract concepts but not as much for concrete concepts are taken to support emotional grounding. On the other hand, results where emotional valence plays a stronger role for specifically *concrete* concepts are taken as evidence against emotional grounding.

To be relevant for the emotional grounding hypothesis, a result needs to include both concrete and abstract words. As an example of a study that does not directly speak to the issue of emotional grounding, consider Ponari *et al.* [66], who conducted a study showing that more emotionally valenced abstract concepts were acquired more easily than neutral abstract concepts. While these results are important, they do not speak to the emotional grounding hypothesis directly because there is no comparison with concrete concepts, i.e. the study does not allow concluding that the facilitatory role of emotionality is specific to abstract concepts. Thus, the following sections (§§2a,b) focus on studies that include both concrete and abstract concepts to allow a comparison with respect to how much they interact with emotional valence.

(a) Evidence for emotional grounding

A key piece of evidence for emotional grounding is the fact that emotional valence ratings are related to concreteness ratings in the manner of an inverted U-shaped pattern. Specifically, relatively more strongly positive as well as relatively more strongly negative concepts are relatively less concrete [17,59]. This pattern has generally been captured with polynomial regression models that include a quadratic emotional valence effect, thereby modelling the fact that more neutral concepts are relatively more concrete, and both positive and negative concepts are more abstract. Newcombe *et al.* [57] also showed that a measure of 'emotional experience'—the relative ease with which a concept evokes emotional experiences—is negatively correlated with concreteness ratings ($r = -0.26$). Similarly, Villani *et al.* [55] demonstrated a positive correlation between abstractness ratings and emotionality ratings for Italian ($r = 0.24$).

A key reaction time study in this field of research was conducted by Kousta *et al.* [16], who unexpectedly found that abstract concepts are processed *faster* than concrete ones when important lexical variables are controlled for. The authors relate this observation to the fact that emotionally valenced words are processed faster than neutral words that are not strongly positive or negative [7,67,68]. This interpretation is supported by an additional analysis which shows that the residual processing speed advantage of abstract concepts is accounted for by entering emotional valence as a covariate [16].

Newcombe *et al.* [57] show that for abstract words, emotional experience was associated with faster and more accurate semantic categorizations. Pauligk *et al.* [4] found that higher positive or negative emotional valence led to lower error rates specifically for abstract but not for concrete words in a lexical decision task. Moffat *et al.* [56] showed that emotional experience facilitated responses to abstract words in a verbal semantic categorization task, but this only happened when stimuli where blocked by emotional experience, thereby drawing attention to the emotional dimension. On top of facilitation effects, Siakaluk *et al.* [58] were also able to demonstrate semantic interference effects of emotional experience for specifically abstract concepts, but again only when stimuli were blocked by emotional experience to make this dimension more salient to participants.[1] A functional magnetic resonance imaging (fMRI) study conducted by Vigliocco *et al.* [59] furthermore found that, only for abstract but not for concrete concepts, emotional valence manipulated haemodynamic activity in the rostral anterior cingulate cortex (rostral ACC), an area that has been implicated in emotion processing [69,70]. Subsequent studies, however, failed to replicate this interaction between valence and concreteness in the rostral ACC [71].

Using age-of-acquisition rating data, Ponari *et al.* [17] showed that abstract words that are very positive or very negative are acquired earlier. For concrete concepts, only positive concepts are acquired early. These results are partially consistent with the emotional grounding hypothesis. Ponari *et al.* [17] also conducted a reaction time study with children aged 6–11, finding that the middle age group in this cohort (8–9-year-olds) showed an effect of emotionality for abstract but not concrete concepts, although this effect was only observed for positive not negative valence. In a similar vein, Lund *et al.* [51] conducted a processing study with three age groups between 5 and 7 years, also finding that the middle age group (here 6-year-olds) showed an effect of emotional valence that was specific to abstract concepts, but only for positive not negative valence. Finally, Kim *et al.* [72] showed that for recognition memory in 7–8-year-old children, valence interacted with abstract concepts, but only for negative words. These acquisition studies are generally interpreted as supporting the idea of emotional grounding, although the fact that only partial effects were obtained (i.e. for either positive or negative valence, not both) deserves more scrutiny. It is possible to construe the existence of partial effects as a disconfirmation of the idea of emotional grounding as originally formulated (e.g. [16]), given that the original claim was predicated on both positive and negative valence.[2]

(b) Evidence against emotional grounding

Yao *et al.* [7] found a facilitatory effect of emotional valence for concrete but *not* for abstract words. This finding has exactly the opposite sign to what was originally reported by Kousta *et al.* [16], although very similar types of variables were entered into the statistical analysis as controls. In addition, Yao *et al.* showed that individual differences in alexithymia (difficulty in identifying and describing emotions) did not modulate the interaction between the abstract/concrete dimension and valence, as would be expected if affective processing were a necessary component for understanding abstract concepts. Additional evidence inconsistent with the emotional grounding hypothesis comes from Kanske & Kotz [73], who found an emotion effect on reaction times only for concrete words, but not for abstract words. Event-related potentials (ERPs) furthermore revealed that concreteness and valence interacted in the late positive component (LPC), an ERP signature that has been linked to mental imagery. Using recordings of facial muscle activity, Künecke *et al.* [74] found that highly valenced words led to increased electrical activity in the corrugator supercilii muscle, a muscle involved in frowning that is highly correlated with stimulus valence [75]. In opposition to the idea of emotional grounding, this valence effect was only observed in response to concrete but not in response to abstract words. In an ERP study, Palazova *et al.* [76] found that emotion-related differences in an early posterior negativity (EPN) arose faster for concrete than for abstract verbs.

To conclude this brief review, the evidence for emotional grounding is clearly mixed: many studies find stronger behavioural effects of emotional valence for concrete concepts, and not abstract concepts. This mixed evidence calls for a deeper investigation of the key results that are claimed to support the idea of emotional grounding. This paper focuses on what Ponari *et al.* [66, p. 1856] have called the 'starting point' for emotional grounding, which is that there is 'a general statistical preponderance of affective information for abstract words' [16, p. 25]. For the current investigation, I take the inverted U-shape (both positively and negatively valenced words are more abstract) as the signature of the emotional grounding hypothesis and assess the extent to which this nonlinear pattern generalizes across languages, concepts and rating scales.

3. Extending emotional grounding beyond English

(a) Rationale

It is problematic to take English and other European majority languages as a vantage point when claims are actually predicated upon the conceptual system writ large [77,78]. A cross-linguistic test of the idea of emotional grounding is especially important because it is known that cultures differ with respect to emotion concepts [79,80], and corpus analyses show that the meanings of emotion-related concepts are not well aligned across cultures [81].

To assess the cross-linguistic generalizability of the emotional grounding hypothesis, the inverted U-shape relationship between emotional valence and concreteness reported for English by Vigliocco *et al.* [59] and Ponari *et al.* [17] will be assessed for the languages presented in table 1, which resulted from an extensive search for cross-linguistic rating studies. Although this dataset includes only three non-Indo-European languages (Indonesian, Mandarin Chinese, Cantonese), considering ten different languages is a considerable improvement vis-à-vis the existing literature. Some studies have investigated the idea of an inverted U-shaped relationship between emotional valence and concreteness for particular languages, but this is the first study, to my knowledge, on this topic to synthesize results from across rating studies. All studies adopt a similar definition of concreteness, including similar instructions given to participants before conducting the rating task.

(b) Methods

The brms package version 2.16.2 was used to fit Bayesian regression models [94,95], and the tidyverse package version 1.3.1 [96] was used for data processing. All analyses were

royalsocietypublishing.org/journal/rstb *Phil. Trans. R. Soc. B* **378**: 20210368

Table 1. Languages and rating datasets considered in this study.

language	*N* words	source
Cantonese	290	Yee [82]
Mandarin Chinese	1100	Yao *et al.* [83]
Croatian	3022	Ćoso *et al.* [84]
Dutch	valence: 4300; concreteness: 30 000	Moors *et al.* [85]; Brysbaert *et al.* [86]
French	valence: 1000; concreteness: 1660	Monnier & Syssau [87]; Bonin *et al.* [88]
German	1000	Kanske & Kotz [89]
Indonesian	1490	Sianipar *et al.* [90]
Italian	1120	Montefinese *et al.* [91]
Polish	4900	Imbir [92]
Spanish	1400	Guasch *et al.* [93]

conducted with R version 4.1.1 [97]. The tidybayes package version 3.0.1 [98] was used for plotting posterior distributions. The patchwork package version 1.1.1 [99] was used for creating multi-plot layouts. Throughout this paper, I use Bayesian regressions for the main analyses, but also report frequentist models when claims specifically relate to existing analyses for which *p*-values were the inference criterion. Data and analysis code are available from the Open Science Framework repository: https://osf.io/8p2an/

In line with standard practice in the analysis of rating data, the individual word (averaged across ratings per participant) is the unit of analysis in the statistical models considered here. To calculate R^2 for the different languages, individual regression models were fitted, for which per-word concreteness ratings were regressed onto two predictors: valence, and valence-squared. Throughout the paper, I use such polynomial regression models. The rsq package version 2.2 [100] was used to calculate partial R^2 to allow looking at how much variance is described by the linear and quadratic effects, respectively.

To generalize across languages, the main Bayesian model considered all rating data, *z*-scored within languages to standardize the different scales, with a random effect for language. The model included by-language varying random slopes for both the linear and the quadratic valence effects, which is needed to support the claim that these effects generalize across languages [101,102]. The model we considered here does not include a random effect for item because (1) datasets have different concepts, many of which do not overlap between the rating studies, and (2) using a random effect for 'item' amounts to assuming translational equivalence between the concepts across languages. As most concepts have non-overlapping glosses, the matching of concepts across languages is hard and laden with assumptions.

For more conservative inferences and to avoid overfitting [103,104], a weakly informative prior was set on all regression slopes for the main Bayesian model: Normal($\mu = 0$, $\sigma = 0.2$). Other than this, I followed the default priors automatically assigned by the brms package. Finally, when considering rating data, it is important to consider that. for some concepts, participants disagree more in their ratings, as reflected in the corresponding standard deviations [65]. This was dealt with by adding standard deviations as regression weights to the linear mixed effects model, a method that has been shown to improve model fits for studies analysing concreteness ratings [63]. These regression weights penalize high-standard-deviation words. Doing this improved the fit of the model (from $R^2 = 0.09$ to $R^2 = 0.13$), which suggests that it is useful to incorporate disagreement between raters into the model.

Markov chain Monte Carlo estimation was executed with four chains and for 8500 iterations (6500 warmup samples discarded). There were no divergent transitions and all chains mixed well ($\hat{R} = 1.0$ for all parameters).

(c) Results

Figure 1 shows scatter plots of concreteness ratings across the emotional valence predictor, with super-imposed linear regression fits (maximum-likelihood point estimate) of the corresponding polynomial regression models. Table 2 summarizes the quadratic coefficients and partial R^2 of the quadratic effect. As can be seen from both the figure and the table, there are negative quadratic effects for almost all languages except for Mandarin Chinese and Cantonese. It is noteworthy that I failed to reproduce the quadratic effect reported in Yao *et al.* [83], i.e. the same data do not yield the inverted U-shaped pattern reported in the original study.

The next analysis considers all languages together in a single linear mixed effects model with random effects (see Methods), allowing generalization across this set of languages. Figure 2 shows the posterior distributions of the linear and quadratic coefficient from this conjoined analysis. As can be seen, the posterior distribution of the quadratic coefficient (posterior mean: −0.22, s.e. = 0.06) is far away from zero, with a 95% credible interval excluding zero, [−0.33, −0.10]. The posterior probability of this effect being of the same sign is very high, $p(\beta < 0) = 0.99$. In contrast to the quadratic effect, the posterior distribution of the linear coefficient (posterior mean: + 0.06, s.e. = 0.07) firmly includes zero; 95% credible interval: [−0.07, +0.19]. The posterior probability of this effect being of the same sign is $p(\beta > 0) = 0.84$. The Bayesian mixed model describes 13% of the variance in concreteness ratings.

Together, these results provide support for the idea that the emotional grounding hypothesis characterizes this set of languages: it appears as if more strongly emotionally valenced concepts—both positive and negative—are relatively more abstract in all of the languages except for Mandarin Chinese and Cantonese. However, this result definitely does not

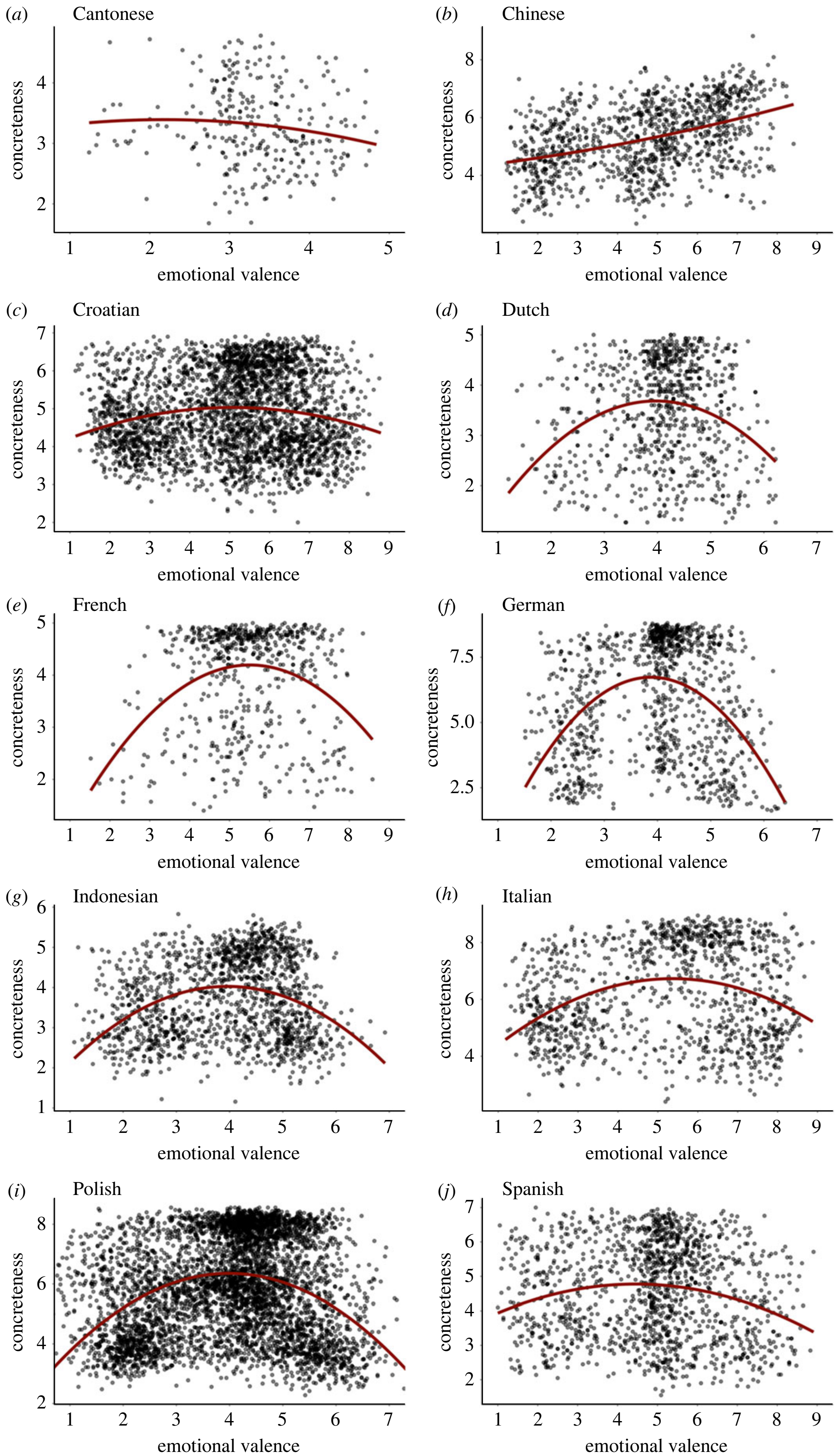

Figure 1. Scatter plots of all words in concreteness × emotional valence space for the respective languages, with super-imposed linear regression fits. (Online version in colour.)

permit any conclusion that emotional grounding is a cross-linguistic universal. Research in linguistic typology generally requires many more languages from a much more diverse sample, including more languages from other language families. Given that rating data are only available for the small set of languages discussed here—all of which are

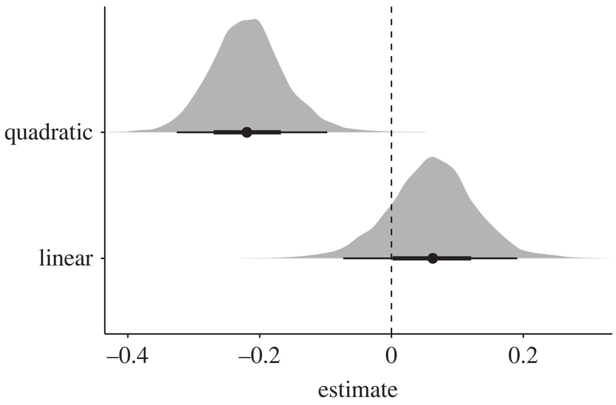

Figure 2. Posterior distributions for the linear and quadratic coefficient of the Bayesian linear mixed effects model; the dashed vertical line shows zero.

Table 2. Quadratic coefficients and standard errors extracted from the corresponding polynomial regressions; these models include regression weights penalizing high-standard-deviation words (cf. [63]).

language	quadratic effect	partial R^2 of quadratic effect
Cantonese	−0.04, s.e. = 0.07	0.0008
Mandarin Chinese	+0.0095, s.e. = 0.01	0.0007
Croatian	−0.06, s.e. = 0.007	0.02
Dutch	−0.30, s.e. = 0.03	0.10
French	−0.17, s.e. = 0.02	0.13
German	−1.06, s.e. = 0.56	0.27
Indonesian	−0.22, s.e. = 0.02	0.09
Italian	−0.14, s.e. = 0.01	0.03
Polish	−0.33, s.e. = 0.01	0.13
Spanish	−0.09, s.e. = 0.01	0.04

associated with large industrialized societies—we simply do not know whether the emotional grounding hypothesis is a genuine universal tendency. In particular, we do not know whether it would also characterize data from minority languages and languages primarily spoken in rural communities, or communities with little contact to industrialized societies. Nevertheless, the fact that the inverted U-shaped emerges as a reliable effect when data are aggregated across languages does suggest limited cross-linguistic generalizability.

4. Do the original data support emotional grounding?

(a) Rationale and approach

Statements about emotional grounding are generally predicated upon *all* abstract concepts, e.g. 'emotion provides grounding for abstract concepts' [17, p. 2]. An issue with the analyses conducted in the last section is that internal variation in abstract concepts is not captured by a regression model that only incorporates a continuous rating scale. As mentioned above, researchers studying abstract concepts have recently begun to emphasize more strongly that abstract concepts are characterized by heterogeneity [36,55,61,62]. A quick look at figure 1 shows that, for some of the languages considered here, clusters of concepts are readily visible to the naked eye. This is problematic for interpreting the quadratic effect in a continuous manner, as small subgroups of words can create quadratic patterns in the average. To demonstrate that this is actually a concern for the emotional grounding hypothesis, it is useful to briefly consider simulated data. In figure 3*a*, 100 emotional valence values were drawn from a uniform distribution, Uniform($a = 1, b = 9$). An additional 100 concreteness values were drawn from Normal($\mu = 3.03, \sigma = 1.04$), for which the mean and standard deviation correspond to the ratings from Brysbaert *et al.* [105]. As is to be expected given how these data have been initialized, a regression model including linear and quadratic emotional valence reveals no 'significant' quadratic effect in this randomly generated data (coefficient of quadratic effect: −0.003, s.e. = 0.02, $p = 0.85$).

Small variations to this basic set-up can create apparent quadratic effects. For example, if we add a small cluster of only 30 concrete neutral words to the existing 100 data points, this will exert a pull on the quadratic coefficient, creating the inverse U-shaped pattern seen in figure 3*b*. Although the cluster is barely visible in the plot, the quadratic relationship would be judged to be 'significant' ($b = -0.04$, s.e. = 0.16, $p = 0.02$). To some extent, this average quadratic trend is real and indeed an accurate reflection of the relationship between emotional valence and concreteness for these data. However, the quadratic trend could also be seen as spurious, given that we know that it does not characterize the whole concreteness rating scale but is instead driven by only a small group of words. The majority of words (those that are also shown in the original figure 3*a*) do not actually follow the quadratic trend that is suggested by the regression model. Clearly, for the simulated data shown in figure 3*b*, the general claim that 'abstract concepts are more emotional' does not apply.

An additional way of creating quadratic patterns is to add 20 negative words and 20 positive words with high abstractness, resulting in figure 3*c*. Again, a quadratic pattern emerges in the average, but we know that it is entirely driven by these two small sets of words. These clusters are barely visible to the naked eye, but they are enough to create an average quadratic effect in the corresponding regression model that would be judged to be 'significant' ($b = -0.04$, s.e. = 0.02, $p = 0.04$).

(b) Applying cluster analysis to the original English data

The simulated data represent a proof-of-concept demonstration of the idea that clusters are a potential problem for the emotional grounding hypothesis. Whether there actually is statistical support for clusters in the rating data is a separate question. To assess the impact of clusters on the emotional grounding hypothesis, figures 4 and 5 reproduce the analyses by Vigliocco *et al.* [59] (using the MRC concreteness ratings [106] and ANEW emotional valence ratings [107]) and Ponari *et al.* [17] (using the Brysbaert *et al.* concreteness ratings [105] and the Warriner *et al.* [108] emotional valence ratings). For both analyses, there were 'significant' quadratic effects (Vigliocco *et al.*: $b = -13.12$, s.e. = 1.14, $p < 0.0001$; Ponari *et al.*: $b = -0.07$, s.e. = 0.004, $p < 0.0001$), consistent with emotional grounding, and consistent with what was reported in the original studies.

The mclust package version 5.4.7 [109] was used to perform cluster analyses using Gaussian mixture models over the two-dimensional space spanned by concreteness and

royalsocietypublishing.org/journal/rstb *Phil. Trans. R. Soc. B* **378**: 20210368

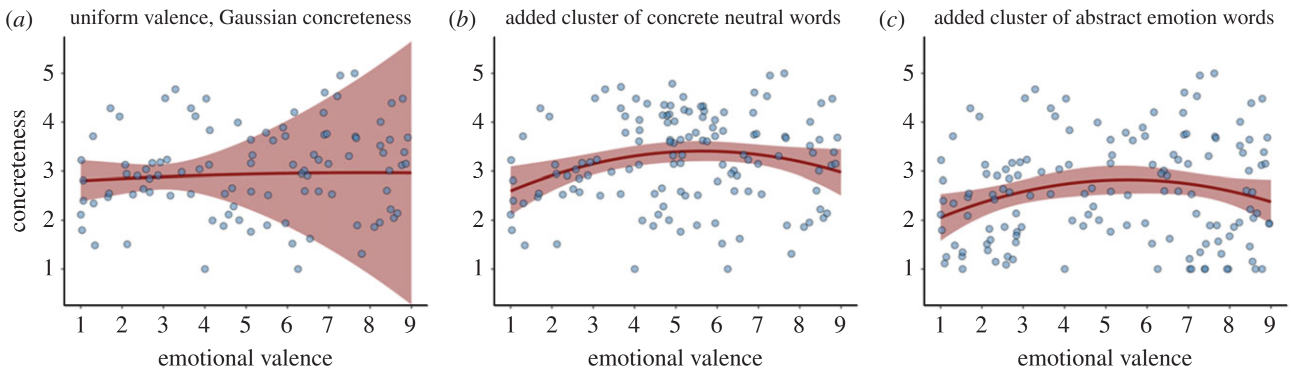

Figure 3. (*a*) A random cluster of words showing no quadratic effect; (*b*) a small group of concrete neutral words has been added to the data from (*a*); (*c*) a small group of abstract emotion words has been added to the data from (*a*). (Online version in colour.)

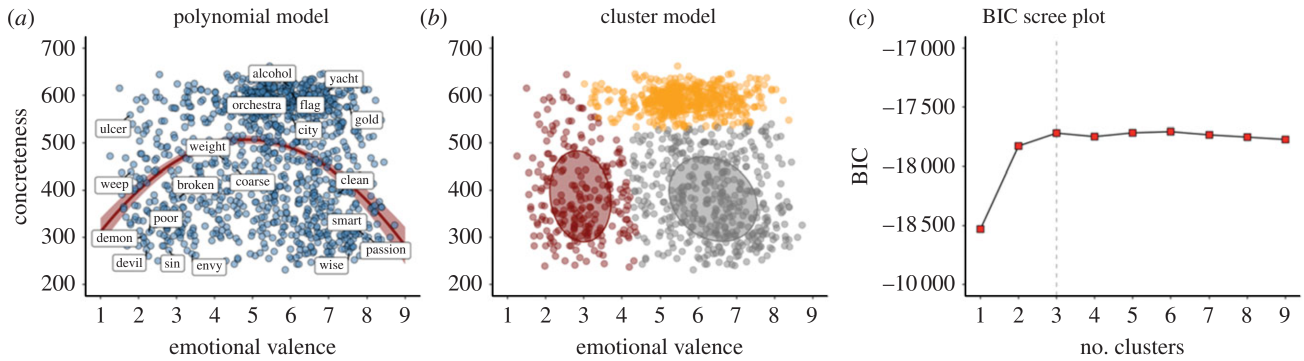

Figure 4. (*a*) Polynomial model for the MRC concreteness ratings [106] and the ANEW emotional valence ratings [107], replicating Vigliocco *et al*. [59]; (*b*) shows the same data with super-imposed Gaussian mixture models for a three-cluster solution, as determined by (*c*) a scree plot of Bayesian information criterion (BIC) against clusters (the mclust package [109] uses a reversed BIC scale). (Online version in colour.)

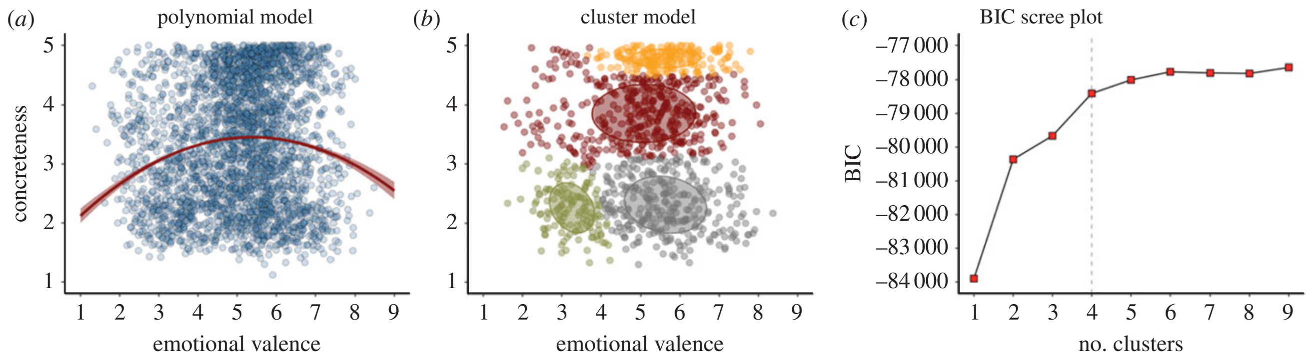

Figure 5. (*a*) Polynomial model for the Brysbaert *et al*. [105] concreteness ratings and the Warriner *et al*. [108] emotional valence ratings, replicating Ponari *et al*. [17]; (*b*) shows the same data with super-imposed Gaussian mixture models for a four-cluster solution, as determined by (*c*) a scree plot of Bayesian information criterion (BIC) against clusters (the mclust package uses a reversed BIC scale). (Online version in colour.)

emotional valence. Scree plots of Bayesian information criterion (BIC) values (mclust uses a reversed BIC scale) were used to determine cluster solutions for each dataset.[3] For the data from Vigliocco *et al*. [59], a three-cluster solution emerged as adequate; for the data from Ponari *et al*. [17], a four-cluster solution emerged as adequate.

When concreteness ratings are regressed onto a three-cluster solution of the Gaussian mixture model for the Vigliocco *et al*. [59] data, this predictor explains considerably more variance (adjusted $R^2 = 0.69$). If clusters and linear and quadratic effects are simultaneously entered into the same regression model, the quadratic effect ceases to be 'significant' ($b = 0.90$, s.e. $= 0.85$, $p = 0.29$). For the data from Ponari *et al*. [17], the variance explained by a four-cluster solution vastly supersedes the variance explained by the quadratic model (adjusted $R^2 = 0.69$). When both polynomials (linear and quadratic effect) and clusters are simultaneously entered into the same regression analysis, the model suggests a quadratic effect in the opposite direction ($b = 0.005$, s.e. $= 0.002$, $p = 0.007$).[4]

The same Gaussian mixture model approach was applied to all cross-linguistic datasets. For every single language except for Cantonese, a three-cluster solution emerged as optimal. Just as was the case with the English data, when concreteness was regressed onto the cluster predictor, this described between 36 and 82% more variance than the corresponding quadratic model, with the exception of Cantonese. When both clusters and polynomials were simultaneously entered into the same regression models, there were 'significant' negative quadratic effects for Dutch, Indonesian,

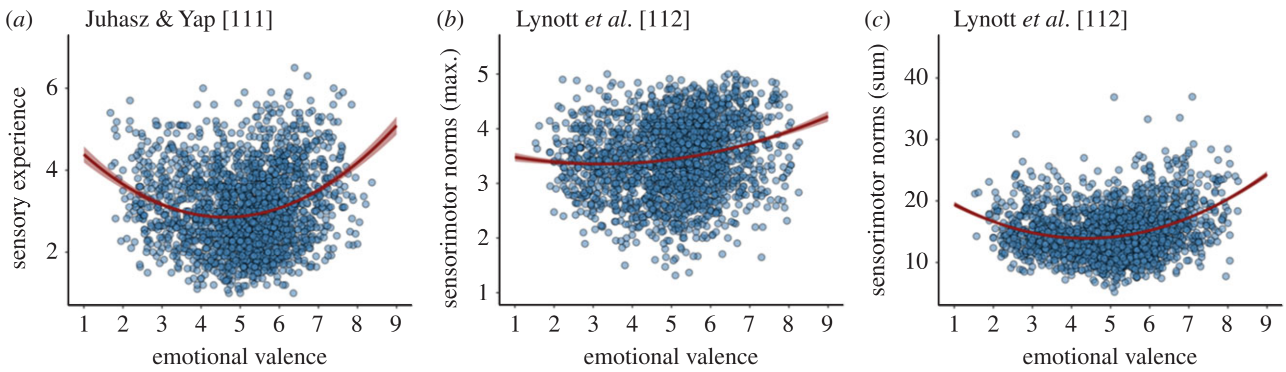

Figure 6. Emotional valence from Warriner *et al.* [108] plotted against sensory experience ratings [111] and the maximum and sum of the Lancaster sensorimotor norms [112], showing a non-inverted U-shaped pattern that is diametrically opposed to the emotional grounding hypothesis. (Online version in colour.)

German, French and Mandarin Chinese, 'significant' *positive* quadratic effects for Croatian and Italian, as well as no 'significant' effects for Spanish, Polish and Cantonese. See the online open science framework repository (https://osf.io/8p2an/) for detailed results for all individual languages.

(c) Effect size considerations

The cross-linguistic data shown in figure 1 as well as the original English data shown in figures 4 and 5 also raise another reason for concern. The intense scatter suggests that the majority of abstract concepts are *not* captured by the quadratic trend: across the whole range of the emotional valence spectrum, we find words of all concreteness levels, a pattern that is particularly striking for the data from Ponari *et al.* [17] shown in figure 5*b*. Standardized effect size measures paint a similar picture. For the data from Vigliocco *et al.* [59], there is a quadratic effect that is associated with 10.5% partial variance. For the data from Ponari *et al.* [17], partial R^2 was even lower, with only 2.8% of the overall variance in concreteness ratings being attributable to the quadratic effect. Together with the visual impression suggested by the scatter plots, the effect sizes show that emotional grounding clearly fails to account for most of the variation in concreteness ratings. The partial R^2 values for the cross-linguistic data shown in table 2 suggest a similarly humbling picture, with 6 out of 10 regression models describing less than 10% of the variance.

5. Triangulating concreteness using different rating scales

So far, I have explored the generalizability of the emotional grounding hypothesis with respect to different languages (§3) and analysis methods (§4). Another important aspect of assessing the generalizability of this hypothesis is the extent to which it depends on using a particular operationalization of concreteness, specifically concreteness ratings. There are many different ways of operationalizing concreteness [20,21,36,110]. Some studies provide direct empirical evidence that other rating scales capture variance in behavioural data better than concreteness ratings [64]. If we take accessibility to the senses as a primary component of the concrete/abstract distinction, scales specifically focused on sensory experience [111] or perceptual strength [53,64,112,113] are viable alternative operationalizations. To see whether the inverted U-shape holds for such other measures, I combined emotional valence ratings from Warriner *et al.* [108] with sensory experience ratings from Juhasz & Yap [111] and the Lancaster sensorimotor norms [112].

Figure 6*a* shows that, for sensory experience ratings from Juhasz & Yap [111], there is a *positive* rather than negative quadratic coefficient ($b = 0.12$, s.e. = 0.007, $p < 0.0001$), i.e. the U-shaped curve is not inverted. The same is the case for the maximum sensorimotor strength association from the Lancaster modality norms [112] ($b = 0.03$, s.e. = 0.003, $p < 0.0001$), as well as for the sum of all sensorimotor ratings from the same data ($b = 0.49$, s.e. = 0.02, $p < 0.0001$).[5] And even though I used the same emotional valence data, all R^2 values for these quadratic trends exceed the effect sizes reported above for the Ponari *et al.* [17] data, ranging from sensorimotor maximum (3%), to sensory experience ratings (6%), to sensorimotor sum (8%). These new results are diametrically opposed to the idea of emotional grounding. It appears that once we move away from concreteness ratings, it is concepts that are *more* concrete that are also more emotionally valenced.

6. Discussion

The results presented here generalize the idea of emotional grounding to a larger set of languages, but they also suggest strong limitations. The following sums up all data-driven results that speak against the emotional grounding of abstract concepts:

— Scatter plots suggest that it is in fact not the case that the quadratic effect expected under emotional grounding characterizes the majority of abstract concepts.
— This is also suggested by the relatively weak effect sizes ($R^2 = 0.028$ in the biggest English dataset).
— A proof-of-concept demonstration with simulated data shows that clusters can create apparent quadratic effects in the kinds of polynomial regression models originally used to support the emotional grounding hypothesis.
— Cluster analyses performed on the original data and the new cross-linguistic data reveal clear subgroups of words that—as suggested by the simulated data—could be driving quadratic trends in the average.
— Cluster models outperform polynomial ones in terms of variance described; at least for English, quadratic effects would cease to be 'significant' once clusters are accounted for.

— Finally, other measures for operationalizing concreteness show quadratic effects with opposite sign where it is concrete concepts rather than abstract concepts that are more emotional.

These are the new data-driven arguments against emotional grounding presented in this paper. These arguments are in line with an issue that has previously been raised against the idea of emotional grounding, which is that many abstract concepts, such as 'even', 'number' and 'proton', do not appear to have strong connections to emotion [41, p. 382; 22, p. 274]. The cluster analysis by Villani *et al*. [55] also provided direct empirical evidence that there are at least some subgroups of abstract concepts that have few ties to emotion. The results presented here are much in line with a view where emotional and affective content could plausibly play a role for the representation of *some* abstract concepts, but, as directly demonstrated here with new analyses, clearly not the majority of abstract concepts. From this perspective, general claims such as 'emotion provides grounding for abstract concepts' [17, p. 2] need to be qualified. Emotion may provide grounding only for a small minority of concepts.

Moreover, we have to remind ourselves of the existence of several behavioural studies that produce effects that directly contradict the idea of emotional grounding, as reviewed by Borghi *et al*. [22, p. 274], and as reviewed in §2 [7,73,74,76]. This includes the fact that at least one key result—the interaction between valence and rostral ACC activation specifically for abstract concepts [59]—has failed to replicate [71]. In addition, as also discussed above, several studies find that emotionality only matters for abstract concepts when this dimension is made salient to participants [56,58], which, as pointed out by Borghi *et al*. [22, p. 280], could be seen as positive evidence that 'emotions did not always play a role' in the processing of abstract concepts. The fact that emotion effects can so easily disappear if the emotional dimension is not made prominent to participants is consistent with the weak effect sizes reported here. This fact is furthermore consistent with the observation that emotional valence effects in lexical processing are generally associated with small effect sizes [68], which suggests that emotional valence may play some role in language comprehension, but not a particularly important one. Finally, as also reviewed in §2, some valence effects are only partial, found for either only positive or only negative valence.

All of these issues (weak effect sizes, effects of opposing sign reported here and in the literature, partial effects for only one end of the valence spectrum, effects that only emerge when the salience of emotions is experimentally increased, failures to replicate key neuro-imaging results) need to be incorporated into theorizing about emotional grounding. To move the debate surrounding abstract concepts forward, future analyses on the topic need to address weak effect sizes more directly, and they need to account for the potential presence of clusters. The design of experiments and the way statistical analyses are conducted in this field need to more actively deal with the heterogeneity of abstract concepts. Moreover, we need to be careful with making general claims about *all* abstract concepts when the available empirical evidence actually suggests that the emotional grounding hypothesis may only work for a subset of concepts, and only if a particular operationalization of concreteness is used. At the present stage, the available evidence does not strongly support the idea that emotion is an important general factor for the grounding of the majority of abstract concepts. Proponents of emotional grounding need to more actively address the existing behavioural effects where it is concrete rather than abstract concepts that show stronger interactions with emotional valence. Without actively resolving the currently mixed body of evidence, it is not clear what the scope of the emotional grounding hypothesis is. Finally, consistent with multiple representation theories of abstract concepts, emotional valence should at most be seen as one additional factor that may potentially be relevant for abstract concepts, among many other factors.

Data accessibility. Data and analysis code are available in the Open Science Framework repository: https://osf.io/g7p9e/ [114].

Conflict of interest declaration. I declare I have no competing interests.

Funding. Bodo Winter was supported by the UKRI Future Leaders Fellowship MR/T040505/1.

Acknowledgements. I thank Anna Borghi, members of the BALLAB and ESLP conferences, Louise Connell, Melvin Yap and Barbara Juhasz for useful comments and suggestions.

Endnotes

[1]As pointed out by Borghi *et al*. ([22], p. 18), the fact that these emotion effects only emerge when the emotional dimension is made salient to participants could also be seen as a challenge to the idea of emotional grounding because it shows that emotionality does *not* matter when no blocking occurs.

[2]Both Ponari *et al*. [17] and Kim *et al*. [72] offer plausible *post hoc* explanations why the children in the respective studies showed only partial effects. However, in the case of Ponari *et al*. [17], it was positive valence that stood out; in the case of Kim *et al*. [72] it was negative valence. A convincing general account as to why different partial effects are observed across studies is necessary and, more specifically, an account that can make predictions prior to having seen the data. It is possible to interpret the observed partial effects as a disconfirmation of the emotional grounding hypothesis as it has been originally formulated, given that the original hypothesis was about both positive and negative valence.

[3]Gaussian mixture models were fitted with the argument 'modelNames = VVV', thereby allowing for the most flexible cluster shapes (ellipsoid that can have any orientation).

[4]It should be pointed out that clusters that are derived from the data in a bottom-up fashion will naturally capture more variation. In response to the concern that the higher R^2 of a model with data-derived clusters is statistically inevitable, it should be highlighted that the very existence of clusters (regardless of their subsequent use in regression models) speaks to limitations of the emotional grounding hypothesis, given that this hypothesis is generally predicated upon all abstract concepts.

[5]For the Lancaster data, the maximum and sum are based on both the perceptual and the motor rating scales. Similar results are obtained if only the perceptual rating scales are used.

References

1. Binder JR, Westbury CF, McKiernan KA, Possing ET, Medler DA. 2005 Distinct brain systems for processing concrete and abstract concepts. *J. Cogn. Neurosci.* **17**, 905–917. (doi:10.1162/0898929054021102)
2. James CT. 1975 The role of semantic information in lexical decisions. *J. Exp. Psychol. Hum. Percept.*

Perform. **1**, 130–136. (doi:10.1037/0096-1523.1.2.130)

3. Kroll JF, Merves JS, Jill S. 1986 Lexical access for concrete and abstract words. *J. Exp. Psychol. Learn. Mem. Cogn.* **12**, 92–107. (doi:10.1037/0278-7393.12.1.92)
4. Pauligk S, Kotz SA, Kanske P. 2019 Differential impact of emotion on semantic processing of abstract and concrete words: ERP and fMRI evidence. *Scient. Rep.* **9**, 14439. (doi:10.1038/s41598-019-50755-3)
5. Rubin DC. 1980 51 properties of 125 words: a unit analysis of verbal behavior. *J. Verbal Learn. Verbal Behav.* **19**, 736–755. (doi:10.1016/S0022-5371(80)90415-6)
6. Whaley CP. 1978 Word—nonword classification time. *J. Verbal Learn. Verbal Behav.* **17**, 143–154. (doi:10.1016/S0022-5371(78)90110-X)
7. Yao B, Keitel A, Bruce G, Scott GG, O'Donnell PJ, Sereno SC. 2018 Differential emotional processing in concrete and abstract words. *J. Exp. Psychol. Learn. Mem. Cogn.* **44**, 1064–1074. (doi:10.1037/xlm0000464)
8. Bottini R, Morucci P, D'Urso A, Collignon O, Crepaldi D. 2021 The concreteness advantage in lexical decision does not depend on perceptual simulations. *J. Exp. Psychol. Gen.* **47**, 685–704. (doi:10.1037/XGE0001090)
9. Allen R, Hulme C. 2006 Speech and language processing mechanisms in verbal serial recall. *J. Mem. Lang.* **55**, 64–88. (doi:10.1016/j.jml.2006.02.002)
10. Schwanenflugel PJ, Stowe RW. 1989 Context availability and the processing of abstract and concrete words in sentences. *Reading Res. Q.* **24**, 114–126. (doi:10.2307/748013)
11. Doest LT, Semin G. 2005 Retrieval contexts and the concreteness effect: dissociations in memory for concrete and abstract words. *Eur. J. Cogn. Psychol.* **17**, 859–881. (doi:10.1080/09541440540000031)
12. Fliessbach K, Weis S, Klaver P, Elger CE, Weber B. 2006 The effect of word concreteness on recognition memory. *Neuroimage* **32**, 1413–1421. (doi:10.1016/j.neuroimage.2006.06.007)
13. Paivio A. 1965 Abstractness, imagery, and meaningfulness in paired-associate learning. *J. Verbal Learn. Verbal Behav.* **4**, 32–38. (doi:10.1016/S0022-5371(65)80064-0)
14. Paivio A, Csapo K. 1969 Concrete image and verbal memory codes. *J. Exp. Psychol.* **80**, 279. (doi:10.1037/h0027273)
15. Bergelson E, Swingley D. 2013 The acquisition of abstract words by young infants. *Cognition* **127**, 391–397. (doi:10.1016/j.cognition.2013.02.011)
16. Kousta S-T, Vigliocco G, Vinson DP, Andrews M, Del Campo E. 2011 The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* **140**, 14–34. (doi:10.1037/a0021446)
17. Ponari M, Norbury CF, Vigliocco G. 2018 Acquisition of abstract concepts is influenced by emotional valence. *Dev. Sci.* **21**, e12549. (doi:10.1111/desc.12549)
18. Schwanenflugel PJ. 1991 Why are abstract concepts hard to understand? In *The psychology of word meanings* (ed. PJ Schwanenflugel), pp. 235–262. Hillsdale, NJ: Lawrence Erlbaum.
19. Löhr G. 2021 What are abstract concepts? On lexical ambiguity and concreteness ratings. *Rev. Philos. Psychol.* **13**, 549–566. (doi:10.1007/s13164-021-00542-9)
20. Dunn J. 2015 Modeling abstractness and metaphoricity. *Metaph. Symbol* **30**, 259–289. (doi:10.1080/10926488.2015.1074801)
21. Barsalou LW, Dutriaux L, Scheepers C. 2018 Moving beyond the distinction between concrete and abstract concepts. *Phil. Trans. R. Soc. B* **373**, 20170144. (doi:10.1098/rstb.2017.0144)
22. Borghi AM, Binkofski F, Castelfranchi C, Cimatti F, Scorolli C, Tummolini L. 2017 The challenge of abstract concepts. *Psychol. Bull.* **143**, 263–292. (doi:10.1037/bul0000089)
23. Mahon BZ, Caramazza A. 2008 A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. (Paris)* **102**, 59–70. (doi:10.1016/j.jphysparis.2008.03.004)
24. Barsalou LW. 1999 Perceptual symbol systems. *Behav. Brain Sci.* **22**, 577–660. (doi:10.1017/S0140525X99002149)
25. Barsalou LW. 2008 Grounded cognition. *Annu. Rev. Psychol.* **59**, 617–645. (doi:10.1146/annurev.psych.59.103006.093639)
26. Glenberg AM, Kaschak MP. 2002 Grounding language in action. *Psychon. Bull. Rev.* **9**, 558–565. (doi:10.3758/BF03196313)
27. Stanfield RA, Zwaan RA. 2001 The effect of implied orientation derived from verbal context on picture recognition. *Psychol. Sci.* **12**, 153–156. (doi:10.1111/1467-9280.00326)
28. Zwaan RA, Stanfield RA, Yaxley RH. 2002 Language comprehenders mentally represent the shapes of objects. *Psychol. Sci.* **13**, 168–171. (doi:10.1111/1467-9280.00430)
29. Pulvermüller F. 2005 Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* **6**, 576–582. (doi:10.1038/nrn1706)
30. Pulvermüller F. 2013 How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn. Sci.* **17**, 458–470. (doi:10.1016/j.tics.2013.06.004)
31. Pecher D. 2018 Curb your embodiment. *Top. Cogn. Sci.* **10**, 501–517. (doi:10.1111/tops.12311)
32. Jamrozik A, McQuire M, Cardillo ER, Chatterjee A. 2016 Metaphor: bridging embodiment to abstraction. *Psychon. Bull. Rev.* **23**, 1080–1089. (doi:10.3758/s13423-015-0861-0)
33. Dove G. 2016 Three symbol ungrounding problems: abstract concepts and the future of embodied cognition. *Psychon. Bull. Rev.* **23**, 1109–1121. (doi:10.3758/s13423-015-0825-4)
34. Crutch S, Troche J, Reilly J, Ridgway G. 2013 Abstract conceptual feature ratings: the role of emotion, magnitude, and other cognitive domains in the organization of abstract conceptual knowledge. *Front. Hum. Neurosci.* **7**, 186. (doi:10.3389/fnhum.2013.00186)
35. Borghi AM, Barca L, Binkofski F, Tummolini L. 2018 Varieties of abstract concepts: development, use and representation in the brain. *Phil. Trans. R. Soc. B* **373**, 20170121. (doi:10.1098/rstb.2017.0121)
36. Borghi AM, Barca L, Binkofski F, Castelfranchi C, Pezzulo G, Tummolini L. 2019 Words as social tools: language, sociality and inner grounding in abstract concepts. *Phys. Life Rev.* **29**, 120–153. (doi:10.1016/j.plrev.2018.12.001)
37. Dove G. 2009 Beyond perceptual symbols: a call for representational pluralism. *Cognition* **110**, 412–431. (doi:10.1016/j.cognition.2008.11.016)
38. Meteyard L, Cuadrado SR, Bahrami B, Vigliocco G. 2012 Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* **48**, 788–804. (doi:10.1016/j.cortex.2010.11.002)
39. Vigliocco G, Meteyard L, Andrews M, Kousta S. 2009 Toward a theory of semantic representation. *Lang. Cogn.* **1**, 219–247. (doi:10.1515/LANGCOG.2009.011)
40. Louwerse MM. 2018 Knowing the meaning of a word by the linguistic and perceptual company it keeps. *Top. Cogn. Sci.* **10**, 573–589. (doi:10.1111/tops.12349)
41. Dove G. 2014 Thinking in words: language as an embodied medium of thought. *Top. Cogn. Sci.* **6**, 371–389. (doi:10.1111/tops.12102)
42. Dove G. 2018 Language as a disruptive technology: abstract concepts, embodiment and the flexible mind. *Phil. Trans. R. Soc. B* **373**, 20170135. (doi:10.1098/rstb.2017.0135)
43. Begg I, Paivio A. 1969 Concreteness and imagery in sentence meaning. *J. Verbal Learn. Verbal Behav.* **8**, 821–827. (doi:10.1016/S0022-5371(69)80049-6)
44. Paivio A. 1971 *Imagery and verbal processes*. New York, NY: Holt, Rinehart and Winston.
45. Paivio A. 1991 Dual coding theory: retrospect and current status. *Can. J. Psychol.* **45**, 255–287. (doi:10.1037/h0084295)
46. Paivio A, Walsh M, Bons T. 1994 Concreteness effects on memory: when and why? *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 1196–1204. (doi:10.1037/0278-7393.20.5.1196)
47. Paivio A. 1986 *Mental representations: a dual coding approach*. Oxford, UK: Oxford University Press.
48. Binder JR, Desai RH, Graves WW, Conant LL. 2009 Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* **19**, 2767–2796. (doi:10.1093/cercor/bhp055)
49. Binney RJ, Hoffman P, Lambon Ralph MA. 2016 Mapping the multiple graded contributions of the anterior temporal lobe representational hub to abstract and social concepts: evidence from distortion-corrected fMRI. *Cereb. Cortex* **26**, 4227–4241. (doi:10.1093/cercor/bhw260)
50. Sabsevitz DS, Medler DA, Seidenberg M, Binder JR. 2005 Modulation of the semantic system by word imageability. *Neuroimage* **27**, 188–200. (doi:10.1016/j.neuroimage.2005.04.012)
51. Lund TC, Sidhu DM, Pexman PM. 2019 Sensitivity to emotion information in children's lexical processing.

Cognition **190**, 61–71. (doi:10.1016/j.cognition.2019.04.017)

52. Recchia G, Jones M. 2012 The semantic richness of abstract concepts. *Front. Hum. Neurosci.* **6**, 315. (doi:10.3389/fnhum.2012.00315)
53. Connell L, Lynott D, Banks B. 2018 Interoception: the forgotten modality in perceptual grounding of abstract and concrete concepts. *Phil. Trans. R. Soc. B* **373**, 20170143. (doi:10.1098/rstb.2017.0143)
54. Barsalou LW, Wiemer-Hastings K. 2005 Situating abstract concepts. In *Grounding cognition: the role of perception and action in memory, language, and thinking* (eds D Pecher, RA Zwaan), pp. 129–163. Cambridge, UK: Cambridge University Press.
55. Villani C, Lugli L, Liuzza MT, Borghi AM. 2019 Varieties of abstract concepts and their multiple dimensions. *Lang. Cogn.* **11**, 403–430. (doi:10.1017/langcog.2019.23)
56. Moffat M, Siakaluk PD, Sidhu DM, Pexman PM. 2015 Situated conceptualization and semantic processing: effects of emotional experience and context availability in semantic categorization and naming tasks. *Psychon. Bull. Rev.* **22**, 408–419. (doi:10.3758/s13423-014-0696-0)
57. Newcombe PI, Campbell C, Siakaluk PD, Pexman PM. 2012 Effects of emotional and sensorimotor knowledge in semantic processing of concrete and abstract nouns. *Front. Hum. Neurosci.* **6**, 275. (doi:10.3389/fnhum.2012.00275)
58. Siakaluk PD, Knol N, Pexman PM. 2014 Effects of emotional experience for abstract words in the Stroop task. *Cogn. Sci.* **38**, 1698–1717. (doi:10.1111/cogs.12137)
59. Vigliocco G, Kousta S-T, Della Rosa PA, Vinson DP, Tettamanti M, Devlin JT, Cappa SF. 2014 The neural representation of abstract words: the role of emotion. *Cereb. Cortex* **24**, 1767–1777. (doi:10.1093/cercor/bht025)
60. Lenci A, Lebani GE, Passaro LC. 2018 The emotions of abstract words: a distributional semantic analysis. *Top. Cogn. Sci.* **10**, 550–572. (doi:10.1111/tops.12335)
61. Muraki EJ, Sidhu DM, Pexman PM. 2020 Heterogenous abstract concepts: is "ponder" different from "dissolve"? *Psychol. Res.* (doi:10.1007/s00426-020-01398-x)
62. Desai RH, Reilly M, van Dam W. 2018 The multifaceted abstract brain. *Phil. Trans. R. Soc. B* **373**, 20170122. (doi:10.1098/rstb.2017.0122)
63. Strik Lievers F, Bolognesi M, Winter B. 2021 The linguistic dimensions of concrete and abstract concepts: lexical category, morphological structure, countability, and etymology. *Cogn. Ling.* **32**, 641–670. (doi:10.1515/cog-2021-0007))
64. Connell L, Lynott D. 2012 Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition* **125**, 452–465. (doi:10.1016/j.cognition.2012.07.010)
65. Pollock L. 2018 Statistical and methodological problems with concreteness and other semantic variables: a list memory experiment case study. *Behav. Res. Methods* **50**, 1198–1216. (doi:10.3758/s13428-017-0938-y)
66. Ponari M, Norbury CF, Vigliocco G. 2020 The role of emotional valence in learning novel abstract concepts. *Dev. Psychol.* **56**, 1855–1865. (doi:10.1037/dev0001091)
67. Kousta S-T, Vinson DP, Vigliocco G. 2009 Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* **112**, 473–481. (doi:10.1016/j.cognition.2009.06.007)
68. Vinson D, Ponari M, Vigliocco G. 2014 How does emotional content affect lexical processing? *Cogn. Emot.* **28**, 737–746. (doi:10.1080/02699931.2013.851068)
69. Etkin A, Egner T, Peraza DM, Kandel ER, Hirsch J. 2006 Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* **51**, 871–882. (doi:10.1016/j.neuron.2006.07.029)
70. Kanske P, Kotz SA. 2011 Emotion speeds up conflict resolution: a new role for the ventral anterior cingulate cortex? *Cereb. Cortex* **21**, 911–919. (doi:10.1093/cercor/bhq157)
71. Skipper LM, Olson IR. 2014 Semantic memory: distinct neural representations for abstractness and valence. *Brain Lang.* **130**, 1–10. (doi:10.1016/j.bandl.2014.01.001)
72. Kim JM, Sidhu DM, Pexman PM. 2020 Effects of emotional valence and concreteness on children's recognition memory. *Front. Psychol.* **11**, 615041. (doi:10.3389/fpsyg.2020.615041)
73. Kanske P, Kotz SA. 2007 Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* **1148**, 138–148. (doi:10.1016/j.brainres.2007.02.044)
74. Künecke J, Sommer W, Schacht A, Palazova M. 2015 Embodied simulation of emotional valence: facial muscle responses to abstract and concrete words. *Psychophysiology* **52**, 1590–1598. (doi:10.1111/psyp.12555)
75. Larsen JT, Norris CJ, Cacioppo JT. 2003 Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology* **40**, 776–785. (doi:10.1111/1469-8986.00078)
76. Palazova M, Sommer W, Schacht A. 2013 Interplay of emotional valence and concreteness in word processing: an event-related potential study with verbs. *Brain Lang.* **125**, 264–271. (doi:10.1016/j.bandl.2013.02.008)
77. Wierzbicka A. 2014 *Imprisoned in English: the hazards of English as a default language*. Oxford, UK: Oxford University Press.
78. Henrich J, Heine SJ, Norenzayan A. 2010 Most people are not WEIRD. *Nature* **466**, 29–29. (doi:10.1038/466029a)
79. Jackson JC, Watts J, Henry TR, List J-M, Forkel R, Mucha PJ, Greenhill SJ, Gray RD, Lindquist KA. 2019 Emotion semantics show both cultural variation and universal structure. *Science* **366**, 1517–1522. (doi:10.1126/science.aaw8160)
80. Kövecses Z. 2005 *Metaphor in culture: universality and variation*. Cambridge, UK: Cambridge University Press.
81. Thompson B, Roberts SG, Lupyan G. 2020 Cultural influences on word meanings revealed through large-scale semantic alignment. *Nat. Hum. Behav.* **4**, 1029–1038. (doi:10.1038/s41562-020-0924-8)
82. Yee LT. 2017 Valence, arousal, familiarity, concreteness, and imageability ratings for 292 two-character Chinese nouns in Cantonese speakers in Hong Kong. *PLoS ONE* **12**, e0174569. (doi:10.1371/journal.pone.0174569)
83. Yao Z, Wu J, Zhang Y, Wang Z. 2017 Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behav. Res.* **49**, 1374–1385. (doi:10.3758/s13428-016-0793-2)
84. Ćoso B, Guasch M, Ferré P, Hinojosa JA. 2019 Affective and concreteness norms for 3,022 Croatian words. *Q. J. Exp. Psychol.* **72**, 2302–2312. (doi:10.1177/1747021819834226)
85. Moors A, De Houwer J, Hermans D, Wanmaker S, van Schie K, Van Harmelen A-L, De Schryver M, De Winne J, Brysbaert M. 2013 Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behav. Res.* **45**, 169–177. (doi:10.3758/s13428-012-0243-8)
86. Brysbaert M, Stevens M, De Deyne S, Voorspoels W, Storms G. 2014 Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychol.* **150**, 80–84. (doi:10.1016/j.actpsy.2014.04.010)
87. Monnier C, Syssau A. 2014 Affective norms for French words (FAN). *Behav. Res.* **46**, 1128–1137. (doi:10.3758/s13428-013-0431-1)
88. Bonin P, Méot A, Bugaiska A. 2018 Concreteness norms for 1,659 French words: relationships with other psycholinguistic variables and word recognition times. *Behav. Res.* **50**, 2366–2387. (doi:10.3758/s13428-018-1014-y)
89. Kanske P, Kotz SA. 2010 Leipzig affective norms for German: a reliability study. *Behav. Res. Methods* **42**, 987–991. (doi:10.3758/BRM.42.4.987)
90. Sianipar A, van Groenestijn P, Dijkstra T. 2016 Affective meaning, concreteness, and subjective frequency norms for Indonesian words. *Front. Psychol.* **7**, 1907. (doi:10.3389/fpsyg.2016.01907)
91. Montefinese M, Ambrosini E, Fairfield B, Mammarella N. 2014 The adaptation of the affective norms for English words (ANEW) for Italian. *Behav. Res. Methods* **46**, 887–903. (doi:10.3758/s13428-013-0405-3)
92. Imbir KK. 2016 Affective norms for 4900 Polish words reload (ANPW_R): assessments for valence, arousal, dominance, origin, significance, concreteness, imageability and, age of acquisition. *Front. Psychol.* **7**, 1081. (doi:10.3389/fpsyg.2016.01081)
93. Guasch M, Ferré P, Fraga I. 2016 Spanish norms for affective and lexico-semantic variables for 1,400 words. *Behav. Res.* **48**, 1358–1369. (doi:10.3758/s13428-015-0684-y)
94. Bürkner P-C. 2017 brms: An R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28. (doi:10.18637/jss.v080.i01)

95. Bürkner P-C. 2018 Advanced Bayesian multilevel modeling with the R package brms. *R J.* **10**, 395–411. (doi:10.32614/RJ-2018-017)
96. Wickham H *et al.* 2019 Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686. (doi:10.21105/joss.01686)
97. R Core Team. 2019 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org.
98. Kay M. 2021 *tidybayes: Tidy data and geoms for Bayesian models. R package version 3.0.1*. (doi:10.5281/zenodo.1308151)
99. Pedersen TL. 2020 *patchwork: The composer of plots. R package version 1.1.1*. See https://CRAN.R-project.org/package=patchwork.
100. Zhang D. 2021 *rsq: R-squared and related measures. R package version 2.2*. See https://cran.r-project.org/package=rsq.
101. Winter B, Grice M. 2021 Independence and generalizability in linguistics. *Linguistics* **59**, 1251–1277. (doi:10.1515/ling-2019-0049)
102. Jaeger TF, Graff P, Croft W, Pontillo D. 2011 Mixed effect models for genetic and areal dependencies in linguistic typology. *Linguist. Typol.* **15**, 281–319. (doi:10.1515/lity.2011.021)
103. Lemoine NP. 2019 Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos* **128**, 912–928. (doi:10.1111/oik.05985)
104. McElreath R. 2020 *Statistical rethinking: a Bayesian course with examples in R and Stan*, 2nd edn. New York, NY: CRC Press.
105. Brysbaert M, Warriner AB, Kuperman V. 2014 Concreteness ratings for 40 thousand generally known English word lemmas. *Behav. Res. Methods* **46**, 904–911. (doi:10.3758/s13428-013-0403-5)
106. Coltheart M. 1981 The MRC psycholinguistic database. *Q. J. Exp. Psychol. A* **33**, 497–505. (doi:10.1080/14640748108400805)
107. Bradley MM, Lang PJ. 1999 *Affective norms for English words (ANEW): instruction manual and affective ratings*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.
108. Warriner AB, Kuperman V, Brysbaert M. 2013 Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* **45**, 1191–1207. (doi:10.3758/s13428-012-0314-x)
109. Scrucca L, Fop M, Murphy TB, Raftery AE. 2016 mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *R J.* **8**, 289–317. (doi:10.32614/RJ-2016-021)
110. Langland-Hassan P, Faries FR, Gatyas M, Dietz A, Richardson MJ. 2021 Assessing abstract thought and its relation to language with a new nonverbal paradigm: evidence from aphasia. *Cognition* **211**, 104622. (doi:10.1016/j.cognition.2021.104622)
111. Juhasz BJ, Yap MJ. 2013 Sensory experience ratings for over 5,000 mono- and disyllabic words. *Behav. Res. Methods* **45**, 160–168. (doi:10.3758/s13428-012-0242-9)
112. Lynott D, Connell L, Brysbaert M, Brand J, Carney J. 2020 The Lancaster Sensorimotor Norms: multidimensional measures of perceptual and action strength for 40,000 English words. *Behav. Res. Methods* **52**, 1271–1291. (doi:10.3758/s13428-019-01316-z)
113. Lynott D, Connell L. 2009 Modality exclusivity norms for 423 object properties. *Behav. Res. Methods* **41**, 558–564. (doi:10.3758/BRM.41.2.558)
114. Winter B. 2022 Emotional grounding of abstract concepts. Open Science Framework repository. (https://osf.io/g7p9e/)