Running head: LOCATING THREE EMOTIONS IN SPACE

Placing joy, surprise and sadness in space. A cross-linguistic study

Fernando Marmolejo-Ramos [1]✉

Juan C. Correa [2]

Gopal Sakarkar [3]

Giang Ngo [4]

Susana Ruiz-Fernández [5]

Natalie Butcher [6]

Yuki Yamada [7]

[1] *Gösta Ekman Laboratory, Department of Psychology, Stockholm University, Sweden*

[2] *Fundación Universitaria Konrad Lorenz, Colombia*

[3] *Department of Computer Applications, G.H. Raisoni College of Engineering, India*

[4] *School of Education, The University of Adelaide, Australia*

[5] *Leibniz Knowledge Media Research Center, Germany*

[6] *School of Social Sciences, Business & Law, Teesside University, United Kingdom*

[7] *Faculty of Arts and Science, Kyushu University, Japan*

Authors' note

Acknowledgments = FM-R thanks Iryna Losyeva and Alexandra Marmolejo-Losyeva (La Patulya) for helping with the data entry, Rosie Gronthos for proofreading the manuscript, Jorge I. Vélez for his advice on statistical graphics, and Kimihiro Noguchi, Petar Milin and Luis Benites for providing feedback on the statistical procedures. SR-F thanks Jennifer Müller and Sergio Cervera for helping with the data collection. FM-R was supported by grants EDU2013-46437-R from the Ministerio de Economía y Competitividad (MINECO) of Spain and 2012V/PUNED/0009 from the UNED. YY was supported by Kyushu University Interdisciplinary Programs in Education and Projects in Research Development (27822) and JSPS KAKENHI (15H05709).

Correspondence concerning this article should be addressed to: Gösta Ekman Laboratory, Department of Psychology, Stockholm University, Frescati Hagväg 9A, SE-106 91, Stockholm. ✉: fernando.marmolejo.ramos@psychology.su.se, 💻: http://sites.google.com/site/fernandomarmolejoramos/

Abstract

The valence-space metaphor posits that emotion concepts map onto vertical space such that positive concepts are in upper locations and negative in lower locations. Whilst previous studies have demonstrated this pattern for positive and negative emotions e.g. '*joy*' and '*sadness*', the spatial location of neutral emotions e.g. '*surprise*' has not been investigated and little is known about the effect of linguistic background. In this study we first characterised the emotions *joy*, *surprise* and *sadness* via ratings of their concreteness, imageability, context availability and valence before examining the allocation of these emotions in vertical space. Participants from six linguistic groups completed either a rating task used to characterise the emotions or a word allocation task to implicitly assess where these emotions are positioned in vertical space. Our findings suggest that, across languages, gender, handedness, and ages, positive emotions are located in upper spatial locations and negative emotions in lower spatial locations. Additionally, we found that the neutral emotional valence of *surprise* is reflected in this emotion being mapped mid-way between upper and lower locations onto the vertical plane. This novel finding indicates that the location of a concept on the vertical plane mimics the concept's degree of emotional valence.

Keywords: emotions; embodiment; spatial cognition; social cognition; metaphorical mapping.

Introduction

Interdisciplinary evidence from robotics (Marocco, Cangelosi, Fischer, & Belpaeme, 2010), neuroscience (Hauk & Pulvermüller, 2011) and cognitive psychology (Bekkering & Neggers, 2002) support the so-called theory of embodied cognition (Barsalou, 2008). This theory argues that the processing of concepts is associated with the activation of perceptual and motor systems (see Barsalou, 2008; Binder & Desai, 2011) and such an association is bidirectional; i.e. the activation of sensorimotor systems affects conceptual processing (e.g. see experiments in Rueschemeyer, Lindemann, van Rooj, van Dam, & Bekkering, 2010) and the activation of concepts affects sensorimotor systems (e.g. see experiment in Glenberg & Kaschak, 2002). The relationship between concepts and sensorimotor systems is considered essential for effective social cognition; a type of cognition used in everyday life situations [1]. That is, for example, our perceptual and motor system can influence our cognitive processes (e.g., judgment, thinking, decision making), just as these processes can influence our physical actions in social contexts (e.g. Wilson, 2002).

[1] As discussed at length by other researchers (Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005), abstract concepts, e.g. emotions, have sensorimotor correlates. Indeed, Holstege (1992) explains how the motor pathways connect to the limbic (i.e., emotion) system. Thus, both abstract and concrete concepts seem to have sensorimotor correlates. There is a two-way interaction between them, which is supported by views of embodied cognition (e.g., Havas, Glenberg & Rinck [2007] show how the activation of sensorimotor systems affect the processing of emotion concepts). It has to be acknowledged, however, that not all concepts are entirely made up of sensorimotor representations only and some concepts, e.g. those referring to mental states, can have semantic properties that lack such representations (see Leshinskaya & Caramazza, 2016). It is likely that the activation of non-sensorimotor or sensorimotor properties of a concept is highly task-, stimuli- and context-dependant (evidence in favour of context in concepts' property activation can be found in Lebois, Wilson-Mendenhall, & Barsalou, 2015).

Based on this theory, Casasanto (2009) proposed the body-specificity hypothesis (BSH). The BSH argues that people implicitly associate positive valenced concepts with the side of their bodily space on which they are more skilful. The experiments by Casasanto (2009) supported this prediction showing that right-handers were more likely than left-handers to associate the right space with positive ideas and the left space with negative ideas, while the opposite holds true for left-handed participants. Accordingly, right- and left-handers tended to link good things such as intelligence, attractiveness, honesty, and happiness more strongly with their dominant side. In employing functional magnetic resonance imaging (fMRI) to compare right- and left- handers' brain activity during motor imagery tasks and action-verb understanding, Casasanto (2011) found that while left-hemisphere motor areas were activated in right-handers, right-hemisphere motor areas were activated in left-handers. This finding lends additional support to the BSH from a neuroscience perspective.

In addition to this, Ansorge and Bohner (2013; see also Ansorge, Khalid, & König, 2013) reported a congruency effect when subjects had to categorize spatial words like *up* as elevated or less elevated (i.e. as high or low in the vertical space), as well as categorize affective words like *happy* as positive or negative. Their results support the assumption that valence-vertical space associations exist in semantic memory, so that faster responses were observed when target words were presented in spatially congruent locations (e.g., *happy* in the upper part of a computer screen). Similarly, Meier and Robinson (2004) found that positive valenced words activated higher areas of visual space, while negative words activated lower areas of visual space (Study 2; see also Xie, Wang, & Chang, 2014) and Sasaki, Yamada and Miura (2015) showed that the emotional valence of images is influenced by motor action towards the upper or lower vertical spatial location (see also Sasaki, Yamada, & Miura, 2016).

To further expand on these previous studies, Marmolejo-Ramos, Elosúa, Yamada, Hamm, and Noguchi (2013) examined whether a dominance of the vertical plane exists over the horizontal plane. Their results supported the predictions of the BSH described above, but also showed that the vertical plane is more salient than the horizontal plane in relation to the allocation of valenced words. That is, while a rating task showed that left-handers rated the word *left* as more positive than *right* and right-handers showed the opposite pattern, a word allocation task showed that positively-valenced words were placed in upper locations whereas negatively-valenced words were placed in lower locations regardless of participants' handedness. Thus, the results lend support to the BSH and also indicate a higher saliency of the vertical plane over the horizontal in the allocation of valenced words (recent evidence as to the saliency of the vertical plane over the horizontal plane is further reported by Damjanovic & Santiago, 2016). Note that Marmolejo-Ramos et al. (2013) reported some differences in the rating task among several linguistic groups (see Figure 1 in their paper) but there were no linguistic differences in the word allocation task.

However, in a recent specialized section devoted to research in embodied cognition (Marmolejo-Ramos & D'Angiulli, 2014), one article reported a study about the effect of linguistic factors on the valence-space metaphor. Marmolejo-Ramos, Montoro, Elosúa, Contreras, and Jiménez-Jiménez (2014) evaluated whether gender and cultural factors have an effect on the mapping of valenced sentences on the vertical space. In the first experiment, Colombian and Spaniards had to recall and report specific personal situations or contexts related to *joy, sadness, surprise, anger, fear,* and *disgust;* i.e. participants recalled and reported situations or contexts in which these emotions occur. Results showed that females expressed more contexts than males, and importantly Colombians reported more contexts than Spaniards. Based on these results, the researchers designed a new spatial–emotional congruency verification task including sentences that recreated the most representative contexts for the emotions of *joy* and *sadness* (e.g. John had a good time with his friends). After reading a sentence, participants had to judge whether a probe word, displayed in either a high or low position on the screen, was congruent or incongruent with the previous sentence. The results showed a mapping between emotions and vertical space induced by sentences recreating representative emotional contexts. This evidence is in line with research (e.g., Schubert, 2005) suggesting that perceptions and judgments of abstract concepts are processed in metaphorical ways by estimating its relative position inside a vertical space.

The emotion words *joy* and *sadness* are exemplars of positive and negative emotions that have been studied in the context of other valenced concepts (see for an example the classic study by Bradley and Lang, 1999). While the words *joy* and *sadness* represent highly positive and highly negative valenced concepts that are readily mapped onto upper and lower locations in space (e.g. Ansorge & Bohner, 2013), it is unknown how emotion words with rather neutral valence would be mapped onto space. An emotion word that seems to have a rather neutral valence (e.g. Reali & Arciniegas, 2015) and whose metaphorical location onto space has not been investigated is that of *surprise. Surprise* is broadly defined as the detection of unexpected situations that challenge a person’s beliefs (Reisenzein, 2009, Reisenzein, Meyer, & Niepel, 2012). It is a peculiar emotion that seems to swing between being negative (e.g., when a person is victim of a robbery) and also positive (e.g., when a person finds his friends at home to celebrate his birthday; see also Macedo, Cardoso, Reisenzein, Lorini, & Castelfranchi, 2009). Also, it has been found that less verbal contexts can be reported for *surprise* compared to emotions such as *joy* and *sadness* (Marmolejo-Ramos et al., 2014). Interestingly, though, this emotion has not been studied in the context of embodiment, therefore the current study aims to do so along with the previously examined emotions; joy and sadness.

The first step before investigating how these three emotions are mapped onto space is to characterise them regarding their level of concreteness (i.e. the degree to which the concept denoted by a word refers to a perceptible entity [Brysbaert, Warriner, & Kuperman, 2014]), imageability (i.e. the ease with which a word gives rise to a sensory mental image of the word [Paivio, Yuille, & Madigan, 1968], context availability (i.e. the ease with which a context can be brought to mind in which the person would feel that emotion [Schwanenflugel & Shoben, 1983]) and valence (i.e. the level of positive-negative emotional state attached to what the emotion concept refers to [see Grühn & Scheibe, 2008]). The first objective of the study was met by having several linguistic groups rate these three emotion words. Having the ratings from several linguistic groups enables us to gain a comprehensive picture of these emotion words with regards to the levels listed above. Although linguistic differences are expected in the rating of words (see Figure 1 in Marmolejo-Ramos et al., 2013), it is hypothesised that across linguistic groups these emotions could have medium-to-low levels of concreteness, and medium-to-high levels of imageability and context availability. As shown in Table 1, such levels are expected based on previous studies in which the average concreteness, imageability and context availability ratings for the words joy, surprise and sadness have been reported (see Altarriba, Bauer & Benvenuto, 1999; Altarriba & Bauer, 2004; Brysbaert, Warriner, & Kuperman, 2014)[2].

[2] In regards to the concreteness dimension, that emotion words might have medium-to-low levels of concreteness is further confirmed by research showing that the more emotionally-laden a word is, the more abstract it is rated (see Kousta et al., 2011). It is important to note that even if emotion concepts are appended to the category of abstract concepts, there can be an abstract-concrete continuum such that some emotion words are more abstract than others (see chapter 1 in Borghi & Binkofski, 2014). That there is a continuum in the abstractness-concreteness spectrum within abstract concepts, mimics the degrees of concreteness (understood as affordances) found in sets of concrete words (see Siakaluk et al., 2008; Xue, Marmolejo-Ramos, & Pei, 2015).

In regards to *surprise*, it most likely exhibits lower context availability than *joy* (and possibly *sadness*) as found by Marmolejo-Ramos et al. (2014; see Tables 1 and 2 in the article). Note that in that study participants generated verbal contexts representing six different emotions, including the three emotions studied herein. These researchers found that *surprise* had the lowest number of verbal contexts (*joy* had the highest number of verbal contexts, followed by *fear* and *sadness*). Thus, it is expected to support such finding via a rating task. It could be speculated that fewer verbal contexts and lower context availability ratings for the concept of *surprise* could be attributed to the neutrality of the concept, which, in turn, may hinder thinking of clear-cut scenarios associated with that given emotion.

Regarding emotional valence, it is expected that *joy* will be rated as highly positive, while *sadness* will be rated as highly negative. This result has also been reported in previous studies (see Table 1). In the ratings reported in Bradley and Lang (1999), *surprise* seems to lean towards positivity (see Table 1). However, based on theoretical accounts arguing that *surprise* is a rather neutral emotion (e.g. Macedo et al., 2009), we expect that the valence ratings will indicate surprise is in fact neutral.

With regard to the levels of concreteness, context availability, imageability and valence of each emotion word some variability due to linguistic differences can be expected (see Evans & Levinson, 2009). This will ultimately be reflected in language effects in all of the 12 rating conditions (i.e. three emotion words [*joy*, *surprise*, and *sadness*] × four word rating dimension [concreteness, context availability, imageability, and valence]).

/// TABLE 1 AROUND HERE ///

The second objective of the study was to investigate the allocation of these three emotions in space via various linguistic groups. Finding that the positive emotion *joy* and the negative emotion *sadness* are placed on upper and lower spatial locations respectively would support the findings of Ansorge and Bohner (2013; see also Ansorge et al., 2013; Meier & Robinson, 2004; Xie et al., 2014; 2015). Indeed, finding that right-handers place the words *joy* and *sadness* towards rightward and leftward spatial locations respectively would lend extra support to the BSH (see Casasanto, 2009; 2011). However, based on the results by Marmolejo-Ramos et al. (2013), the distance between *joy* and *sadness* on the horizontal plane (i.e., BSH) is expected to not be significant; rather, it is hypothesised a significant difference between *joy* and *sadness* on the vertical plane exclusively [3]. These findings would then lend support to evidence suggesting a saliency of the vertical plane over the horizontal plane (see Figure 2F in Marmolejo-Ramos et al., 2013). Finding that *surprise* is located half-way between the vertical locations of *joy* and *sadness* would show for the first time that *surprise*'s emotional valence is mapped onto space. Specifically, we expect to find that given the neutral valence of *surprise*, this word would be mapped onto a vertical location near the mid-point (i.e. placed between *joy* and *sadness*). The non-linguistic differences originally reported by Marmolejo-Ramos et al. (2013) in the allocation of valenced words onto space suggest that there could be minimal chances of finding language effects in the allocation of these words.

Methods

Participants

University undergraduate students and members of the community from six different linguistic backgrounds (i.e. English, Hindi, Japanese, Spanish, Vietnamese and German) voluntarily participated in the rating (n=325) and the word allocation (n=362) tasks. The

[3] It could be argued that the valence-space metaphor could ensue in the horizontal plane when the vertical plane is being controlled for. However, a recent study in which the valence-space metaphor is tested independently in the horizontal and the vertical plane, i.e. one of the planes is being controlled for, showed such mapping occurs only in the vertical plane (Xie et al., 2015).

experimental protocol was approved by the ethics committees of the institutions involved in the studies. Participants gave written informed consent in order to abide by the principles of the Declaration of Helsinki. Table 2 reports demographic and descriptive statistic information of the participants (participants whose responses reflected a lack of understanding of the instructions, were illegible, or were incomplete were discarded. Also, participants with incomplete demographic data, e.g. no information about gender, handedness, age or language, were not included in the analyses).

/// TABLE 2 AROUND HERE ///

Materials

The three emotion words *joy*, *surprise* and *sadness* were used in the rating study. The ratings were performed via a simple paper-based task (see Figure 1A). The word location task also consisted of a paper-based task (see Figure 1B).

/// FIGURE 1 AROUND HERE ///

Procedure

Rating task

Participants were asked to rate the three emotions on the following dimensions: concreteness, imageability, context availability and valence. The ratings were made by placing a mark (e.g. via a pen or a pencil) on 10cm horizontal lines; one line for each attribute. On the left end, the scales were labelled as ‘highly abstract’ (concreteness scale), ‘hard to imagine’ (imageability scale), ‘hard to think of a context’ (context availability scale) and ‘highly negative’ (valence scale). On the right end, the scales were labelled as ‘highly concrete’ (concreteness scale), ‘easy to imagine’ (imageability scale), ‘easy to think of a context’ (context availability scale) and ‘highly positive’ (valence scale). The three words were presented to participants for rating in a random order; however, the order of each rating (concreteness, imageability, context availability and valence) for each word was given in a fixed order (see Figure 1A).

Word allocation task

Participants were asked to locate three symbols representing the words *joy*, *surprise* and *sadness* on a 10 cm^2 gridded square (this grid resembles that used in Experiment 2 by Marmolejo-Ramos et al., 2013). A triangle represented *joy*, a square represented *surprise* and a circle represented *sadness* and this matching was used for all participants (see Appendix for supplementary results that reflect the counterbalanced emotion/symbol combinations). The instructions read: "assuming the words *joy*, *surprise* and *sadness* were symbols to be placed in the following square, where would you put them?" Participants were also instructed that each symbol should occupy only one square within the grid, each symbol should occupy different squares in the grid, and each symbol should be drawn only once (see Figure 1B). There were not time restrictions to complete this task.

Design and analyses

The data in both tasks were analysed via high breakdown and high efficiency robust linear regression modelling (see Yohai, 1987) via the 'lmRob' function in the 'robust' R package. For the rating study, the independent variables were participant, i.e. all participants in rating study (P), language, i.e. the six languages studied (L), gender, i.e. males and females (G), handedness, i.e. right- and left-handers (H), age, i.e. the ages of the participants in the rating study (A), word, i.e. *joy*, *surprise* and *sadness* (W) and word dimension, i.e. concreteness, imageability, context availability and valence (D). These factors were hierarchically entered in this order and the dependent variable was the rating values.

For the word allocation study, the independent variables were participant, i.e. all participants in word allocation study (P), language, i.e. the six languages studied (L), gender, i.e. males and females (G), handedness, i.e. right- and left-handers (H), age, i.e. the ages of the participants in the word allocation study (A), and word, i.e. *joy*, *surprise* and *sadness* (W). These factors were entered in this order for the location values obtained in the X and Y axes; i.e. the two dependent variables in the word allocation study. The variables W, H and L were central to this study and added to the model based on previous research showing that they play a part in the mapping of words onto space (see Marmolejo-Ramos et al., 2013; 2014). While the variable D is specific to the rating task, the variables P and A were peripheral to this study and were included to account for their potential effects on the dependent variables. Some of the estimates of the beta weights of the levels of the independent variables (*β*-values) and their associated *t* and *p* values were reported to illustrate their influence on the model. For each hierarchical model, the variability accounted for was estimated as adjusted $R^2 \cdot 100$. The models' fits were compared via ANOVA and robustified *F*-tests (F_r).

Average values and associated measures of deviation were estimated via the median (*Mdn*) and median absolute deviation (*MAD*), respectively. The formula $\pm 1.58 \cdot \left(\frac{IQR}{\sqrt{n}}\right)$, where *IQR* = interquartile range and *n* = sample size, was used to generate 95% CI around the medians for assessing equality of medians at approximately 5% significance level (see McGill, Tukey, & Larsen, 1978). Based on the results of the robust ANOVA model comparison, pairwise comparisons were examined via the degree of CIs overlap between groups of interest (e.g. within levels of a variable or between variables). Non-overlapping CIs were taken as evidence of significant difference between the groups' medians (see Cumming & Finch, 2005; Cumming, 2012). However, when there was some degree of overlap between two or more dependent groups, the Agresti-Pendergast ANOVA test (F_{AP}) was used via the R function 'apanova' (see Wilcox, 2005). The *p* values of multiple comparisons were adjusted via the false discovery rate method, p_{FDR} (Benjamini, & Hochberg, 1995). Pairwise comparisons between two or more independent groups were performed via the Cucconi permutation test, *MC* (Marozzi, 2012; 2014).

Results

The rating results suggested no differences among the three emotion words regarding their concreteness levels. However, *joy* received higher context availability ratings than *surprise* and the three words differed in terms of imageability ratings; i.e. *joy* > *surprise* > *sadness*. Central to this study was the finding that in terms of valence *joy* was rated higher than *sadness* and *surprise*'s average ratings fell between the other two words.

Rating task

Only the models P, P + L + G and P + L + G + H did not have significant *t* and *p* values associated with the β-values. The other models had significant β-values (e.g., in the P + L model: $\beta_{Hindi} = -1.86$ ($t=-6.65$, $p<.001$), in the P + L + G + H + A model: $\beta_{age} = -0.03$ ($t=-2.88$, $p<.01$), in the P + L + G + H + A + W model: $\beta_{sadness} = -1.78$ ($t=-17.11$, $p<.001$), and in the in the P + L + G + H + A + W + D model: $\beta_{context} = 1.49$ ($t=12.42$, $p<.001$)). The variability accounted for by each model was 1.02% (P), 4.57% (P + L), 4.63% (P + L + G), 4.66% (P + L + G + H), 4.82% (P + L + G + H + A), 10.78% (P + L + G + H + A + W), and 18.41% (P + L + G + H + A + W + D). A comparison of the models further suggested that there was an improvement of the fitness of the hierarchical models to the rating data when P, L, and A were added; $F_r = 40.90$, $p<.001$, $F_r = 22.49$, $p<.001$ and $F_r = 7.03$, $p=.006$, respectively. However, the largest improvement occurred when W and D were finally added to the model; $F_r = 111.45$, $p<.001$ and $F_r = 104.77$, $p<.001$, respectively.

The model P was significant in that there were differences in the ratings across participants. For example, whereas a participant in the English sample had a median rating of 3.95 (95% CI [3.15, 4.74]), a participant in the Vietnamese sample had a median rating of 7.7 (95% CI [4.89, 10.50]). Language had an effect on the ratings, which was due to median ratings differing across linguistic groups. For example, while the median rating in the Hindi sample was 5.4 (95% CI [5.18, 5.61]), the median rating in the Japanese sample was 6.5 (95% CI [6.26, 6.73]). The effect of age on the ratings was graphically explored via a scatterplot with linear and smooth fit lines and a correlation test. The results indicated a near-significant positive correlation (r_τ = .02, z=1.87, p=.06) such that, for example, the median rating of participants aged 17 to 25 was 6.7 (95% CI [6.49, 6.90]) and the median rating of participants aged 30 to 35 was 7.95 (95% CI [6.70, 9.19]).

The effect of word type (W) was substantiated by the non-overlap between the confidence intervals around the median ratings for the words *joy*, *surprise* and *sadness*; Mdn_{joy} = 7.6 (95% CI [7.42, 7.77]), $Mdn_{surprise}$ = 6.2 (95% CI [6.059, 6.34]), and $Mdn_{sadness}$ = 5.8 (95% CI [5.54, 6.054]) [4]. In the case of the factor word dimension (D), while the average ratings in the context and imageability dimensions did not differ ($Mdn_{context}$ = 7.4 (95% CI [7.20, 7.59]), $Mdn_{imageability}$ = 7.4 (95% CI [7.24, 7.55])), the average ratings in the concreteness and valence dimensions did ($Mdn_{concreteness}$ = 5.7 (95% CI [5.45, 5.94]), $Mdn_{valence}$ = 5.1 (95% CI [4.80, 5.39])). Also, the ratings for the words in the context and imageability dimensions were higher than the ratings for the words in the concreteness and valence dimensions ($Mdn_{context + imageability}$ = 7.4 (95% CI [7.27, 7.52]) and $Mdn_{concreteness + valence}$ = 5.2 (95% CI [5.02, 5.03]).

[4] For clarity, note that that these values are at a group level (not individual level) and are averaged across the four rating dimensions for each word.

Given the significant effects of W and D on the ratings, their relationship was analysed. Figure 2A shows the ratings of the three words according to the dimension in which they were evaluated. In the concreteness dimension, the median ratings of *joy* (*Mdn* = 5.7 (95%CI [5.27, 6.12])), *sadness* (*Mdn* = 5.2 (95%CI [5.54, 6.25])) and *surprise* (*Mdn* = 5.9 (95%CI [4.73, 5.66])) did not differ (F_{AP} (2, 648) = 1.26, p=0.28). In the context dimension, there were differences between groups (F_{AP} (2, 648) = 4.69, p=0.009) due to the median rating of *joy* (*Mdn* = 7.6 (95%CI [7.27, 7.92])) differing from that of *surprise* (*Mdn* = 7.2 (95%CI [6.96, 7.63])) (F_{AP} (1, 324) = 8.68, p_{FDR}=0.01). Other pairwise comparisons in this dimension, and that involved the word *sadness* (*Mdn* = 7.3 (95%CI [6.88, 7.51])), were not significant (all $p_{FDR} > .05$). There were also differences between *joy* (*Mdn* = 7.8 (95%CI [7.58, 8.01])), *sadness* (*Mdn* = 7.5 (95%CI [7.18, 7.81])) and *surprise* (*Mdn* = 7 (95%CI [6.70, 7.29])) in the imageability dimension (F_{AP} (2, 648) = 14.13, p<.001) due to all pairwise comparisons being significant (all $p_{FDR} < .05$). The non-overlap between the 95% CIs of *joy* (*Mdn* = 9.1 (95%CI [8.89, 9.30])), *sadness* (*Mdn* = 1.65 (95%CI [1.40, 1.89])), and *surprise* (*Mdn* = 5.1 (95%CI [4.98, 5.21])) in the valence dimension indicates the average ratings between these groups differed significantly.

Effects of covariates on the ratings of each emotion word

Emotion word JOY: Analyses of the effects of the covariates Participant (P), Language (L), Gender (G), Handedness (H), and Age (A) on the four types of ratings, revealed an effect of P (i.e. P model) on the context availability (CA), imageability (I) and valence (V) ratings of *joy* (CA: F_r = 15.67, p=5.45e^{-05}; I: F_r = 5.90, p=.01; V: F_r = 16.59, p=3.30e^{-05}). There was also an effect of L (i.e. P + L model) on the CA and V ratings of *joy* (CA: F_r = 12.74, p=.03; V: F_r = 19.03, p=.003). All the other models were not significant; p>.05.

Emotion word SURPRISE: Analyses of the effects of the covariates P, L, G, H, and A on the four types of ratings, revealed an effect of P on the CA and I ratings of *surprise* (CA: $F_r = 4.16$, p=.03; I: $F_r = 15.58$, $p=5.74e^{-05}$). There was also an effect of A (i.e. P + L + G + H + A model) on the V ratings of *surprise* ($F_r = 10.35$, p=.001; a Kendall's tau test did not support this effect: τ=.005, p=.89). All the other models were not significant; p>.05.

Emotion word SADNESS: Analyses of the effects of covariates P, L, G, H, and A on the four types of ratings, revealed an effect of P on the concreteness (C), CA, I, and V ratings of *sadness* (C: $F_r = 13.04$, p<.001; CA: $F_r = 29.77$, $p=2.68e^{-08}$; I: $F_r = 26.10$, $p=1.92e^{-07}$; V: $F_r = 29.96$, $p=2.43e^{-08}$). There was also an effect of A (i.e. P + L + G + H + A model) on the C ratings of *surprise* ($F_r = 4.30$, p=.03; τ=.09, p=.01), an effect of L (i.e. P + L model) on the CA ratings ($F_r = 18.69$, p=.003), and an effect of G (i.e. P + L + G model) on the I ratings ($F_r = 4.39$, p=.03; a Cucconi test did not support this effect: *MC*=1.45, p=.23). All the other models were not significant; p>.05.

Word allocation task

The results showed that while no one factor had effects on the X axis data, in the case of the Y axis, regardless of language, gender, handedness and age, *joy* was located in upper spatial locations and *sadness* in lower spatial locations. The neutral emotional concept of *surprise* was located mid-way between joy and sadness. In regard to the language factor, results were in line with those reported by Marmolejo-Ramos et al. (2013) in that there were some differences among linguistic groups in the rating task but none in the word allocation task.

Robust linear regression on the X axis data

In none of the models the t values associated with the β-values were significant (all $p > .05$). The variability accounted for by each model was 0.02% (P), 0.23% (P + L), 0.28% (P + L + G), 0.45% (P + L + G + H), 0.45% (P + L + G + H + A), and 0.66% (P + L + G + H + A + W). A comparison of the models further suggested no improvement of the fitness of the hierarchical models to the *X* axis data; P model: $F_r = .17$, p=.66, P + L model: $F_r = .34$, p=.99; P + L + G model: $F_r = .44$, p=.49; P + L + G + H model: $F_r = 1.40$, p=.22; P + L + G + H + A model: $F_r = .01$, p=.88, and P + L + G + H + A + W model: $F_r = .54$, p=.90.

The overlap between the confidence intervals for the words when located in the *X* axis suggests they are not positioned differently on the horizontal plane (see Figure 2B). Indeed, although there was variability in the location of the words ($MAD_{joy} = 5.93$, $MAD_{surprise} = 5.93$, and $MAD_{sadness} = 8.89$), the median location for the three words was -1 [5].

Effects of covariates on the horizontal position of each emotion word

Analyses of the effects of the covariates Participant (P), Language (L), Gender (G), Handedness (H), and Age (A) on the X values (e.g. effects of those covariates on the values in the X axis when the word was *joy*) showed that there were nonsignificant results in the X axis (p>.05 in all models for each of the three words).

[5] Even if the medians of the words had aligned towards the left or the right of the square, what matters is that they are aligned; that is, that their median locations in the *X* axis do not differ. If there had been found that, for example, *joy* were around 8, *surprise* were around 0 and *sadness* around -7, then the robust linear modelling should have shown significant effects from any of the variables (e.g. handedness) on the analyses of the *X* axis data. However this did not happen.

Robust linear regression on the Y axis data

The same analysis described above for the data in the *X* axis was performed for the data in the *Y* axis. Only in the last model the *t* values associated with the *β*-values were significant; e.g., $\beta_{surprise}$ = -2.67 (*t*=-6.66, *p*<.001), and $\beta_{sadness}$ = -12.14 (*t*=-29.77, *p*<.001). The variability accounted for by each hierarchical model was .01% (P), 0.26% (P + L), 0.28% (P + L + G), 0.32% (P + L + G + H), 0.37% (P + L + G + H + A), and 49.88% (P + L + G + H + A + W). A comparison of the models suggested an improvement of the fitness of the hierarchical models to the *Y* axis data only when the predictor W was added; P model: F_r = .19, *p*=.66, P + L model: F_r = .40, *p*=.99; P + L + G model: F_r = .18, *p*=.66; P + L + G + H model: F_r = .29, *p*=.58; P + L + G + H + A model: F_r = .46, *p*=.49, and P + L + G + H + A + W model: F_r = 373.43, *p*<.001.

The non-overlap between the confidence intervals for the words when located in the *Y* axis suggests they are positioned differently on the vertical plane (see Figure 2B). There was some variability in the location of the words (MAD_{joy} = 2.96, $MAD_{surprise}$ = 4.44, and $MAD_{sadness}$ = 4.44) and they had notably different locations on the *Y* axis. Specifically, while *joy* was located in the upper end of the square (Mdn_{joy} = 7 (95% CI [6.46, 7.53])), *sadness* was positioned on the lower end of the square ($Mdn_{sadness}$ = -7 (95% CI [-7.58, -6.41])); and *surprise* was placed in between the other two words ($Mdn_{surprise}$ = 3 (95% CI [2.58, 3.41])).

Effects of covariates on the vertical position of each emotion word

There was an effect of P in the cases of *joy* and *sadness* only (*joy*: P model: F_r = 2.03, *p*=.14. *sadness*: P model: F_r = 16.46, p=3.54e^{-05}) such that some participants allocated these words more upward/downward than others (all other models in *joy* and *sadness* had *p*>.05). There was an effect of H in the case of *surprise* only (P + L + G + H model: F_r = 4.25, *p*=.03; a Cucconi test confirmed this difference: *MC* = 3.32, *p*=.03) such that right handers allocated this word higher (*Mdn*=3, (95% CI [2.46, 3.53])) than left handers (*Mdn*=2, (95% CI [0.58, 3.41])). All the other models in *surprise* had p>.05 (See Appendix for supplementary results).

/// FIGURE 2 AROUND HERE ///

Discussion and conclusions

The aim of the rating task was to characterise the words under scrutiny in their concreteness, context availability, imageability, and valence dimensions. The word allocation task aimed to determine the allocation of these three emotions in space by various linguistic groups. Overall, the results suggest that the valence of the emotion words *joy, surprise* and *sadness* (as indicated on the valence dimension in the rating task) are metaphorically mapped onto the vertical plane such that *joy* is located in upper locations, *sadness* is located in lower locations and *surprise* is located mid-way between the other two words (word allocation task).

The results of the rating study agree with previous research in which the concreteness, imageability, context availability, and valence of the words *joy, sadness* and *surprise* have been assessed (see Table 1 and Figure 2A); however, the present results add novel details. It was found that the three words have similar levels of concreteness and are rated as mildly concrete. Although the results showed that overall the three words have medium-to-high levels of imageability, as previous studies have indicated, it was further found that *joy* is more imageable than *sadness* and *sadness* is more imageable than *surprise*. Additionally, the finding that *joy* rated higher than *surprise* in regards to context availability is in line with Marmolejo-Ramos et al. (2014; Tables 1 and 2) in which participants generated less emotional contexts for *surprise* than *joy*. The present results thus corroborate the findings of these authors via a rating task. Finally, in agreement with past research, *joy* was rated as more positive than *sadness* and *surprise* was rated mid-way between the other two emotions. However, the median valence rating of *surprise* (*Mdn* = 5.1 (95% CI [4.98, 5.21])) indicates this word is regarded as neither positive nor negative. This is a novel finding since it empirically demonstrates that *surprise* is a rather neutral emotion concept. It is interesting to note that we found an effect of language in the rating task, but such a factor did not mediate the word allocation task (see below).

The results of the word allocation study confirm that highly positive emotions such as *joy* are mapped onto upper spatial locations, while highly negative emotions such as *sadness* are mapped onto lower spatial locations. This finding is in keeping with research suggesting a metaphorical association between emotion stimuli and the vertical spatial axis (e.g. Ansorge & Bohner, 2013, Ansorge et al., 2013; Damjanovic & Santiago, 2016; Marmolejo-Ramos et al., 2014; Meier & Robinson, 2004; Sasaki et al., 2015; 2016; Xie et al., 2014; 2015). Indeed, the average location of the words on the horizontal axis were no different and handedness had no effect lends extra support to the idea that the vertical plane is more prominent than the horizontal plane for the mapping of emotions onto space as originally suggested by Marmolejo-Ramos et al. (2013). Interestingly, while in the rating task the language and age variables had an influence on the words' ratings, this was not the case in the word allocation task. As shown in Figure 1 in the study conducted by Marmolejo-Ramos et al. (2013), the average ratings of words tend to vary across linguistic groups and as shown by Bird, Franklin and Howard (2001), age of acquisition can correlate with, for instance, the imageability ratings of words. Thus, concluding that language and age have an effect on the ratings of emotion words is not surprising (see for example Evans & Levinson [2009] arguments regarding linguistic diversity). However, in the word allocation task these factors, along with the factors gender and handedness, did not have any effect. The results of the word allocation task hence suggest that, regardless of language, gender, handedness and age, positive words are located in upper spatial areas and negative words are located in lower spatial areas. This result corroborates the findings from Marmolejo-Ramos et al. (2013).

The novel finding is that *surprise* was located mid-way between *sadness* and *joy* in the vertical axis. Although the median location of *surprise* on the vertical axis was not exactly zero, it was located rather close to it (*Mdn* = 3 (95% CI [2.58, 3.41])). Numerically speaking, the exact mid-way location in the vertical axis between where *joy* and *sadness* were located is zero and the exact mid-way location between zero and where *joy* was located is 3.5 (see Figure 2B). Thus, it could be said that a location above 3.5 should be an indication of the word leaning towards positivity, while a value on the *Y* axis below 3.5 should be an indication of the word leaning towards neutrality. Given that the upper arm of the CI around the median rating of *surprise* did not cover 3.5, it is then reasonable to assert that this emotion tends to be located mid-way between *joy* and *sadness* in the vertical spatial plane. This result thus provides further evidence that the neutral emotional valence of *surprise* (as found in the rating task) is reflected in this emotion being mapped mid-way between upper and lower locations onto the vertical plane.

Why is vertical space so salient? It has been argued that locations on the horizontal plane (i.e. left and right) are less salient than locations on the vertical plane (i.e. up and down) since people tend to confuse East-West more than North-South (see Mark & Frank, 1989, as cited in Marmolejo-Ramos et al., 2013). Locations on the horizontal plane are less noticeable as it is equally easy to look left or right. Locations on the vertical plane, on the other hand, are clear in that locations above eye level are immediately observable and therefore more likely to be preferred (i.e. likely to be associated with positive valence) than locations below eye level (see also Freeman, 1975, as cited in Marmolejo-Ramos et al., 2013; see also studies on locatives and comparatives by Clark, Carpenter & Just, 1973). It is thus likely that a mapping of positive valenced concepts (concepts that refer to events, objects and people) onto upper spatial locations is strongly influenced by bodily configuration and experience rather than language, which labels such experiences.

Note that all studies on the valence-space metaphor focus on mapping of the opposite ends of the affective continuum of a concept (e.g. positive emotions vs negative emotions) onto the opposite ends of the vertical plane (e.g. high spatial location vs low spatial location). The results have consistently shown that high spatial locations are associated with positivity and low spatial locations negativity (see Clark et al, 1973, and other references cited herein). No previous studies have investigated the location on the vertical plane of neutrally-valenced concepts. Our study is the first to show that such concepts, exemplified here with the case of *surprise*, are associated with the mid-point (between *joy* and *sadness*) in the vertical plane.

It is worth noting that focused analyses showed there were no language effects on the allocation of the three words in the X and Y axis in the first WAT task but there was a language effect on the allocation of *joy* in the X axis and the allocation of *sadness* in the Y axis in the second WAT task (see Appendix). This finding can be due to simple linguistic variability (see Evans & Levinson, 2009). Interestingly, no covariate had an effect on the allocation of *surprise* in the vertical and horizontal plane. This suggests that while there could be some degree of variability across languages as to the allocation of *joy* and *sadness* in 2D space, there seems to be less variability as to the spatial location of *surprise*. In other words, *surprise* seems to be zeroed in a specific vertical and horizontal coordinate.

This novel result indicates that the location of a concept on the vertical plane mimics the concept's degree of emotional valence regardless of linguistic background. Indeed, it could be entertained that the location of any stimulus on the vertical plane should mimic the stimulus' degree of emotional valence. That is, the more positively valenced the stimulus, the higher in vertical space it would be located; likewise, the more negatively valenced the stimulus, the lower it would be located. By the same token, a stimulus that is neither too positive nor too negative would tend to be located towards the middle in the vertical plane, as *surprise* was found to be here. A recent study by Sasaki et al. (2015) could be modified to verify this claim. Sasaki et al. (2015) had participants evaluate emotional images. Before evaluation responses were made, the participants had to swipe the display upward or downward, and then they made an evaluation of the image's valence. Surprisingly, when participants swiped upward before the evaluation, a more positive evaluation was given to images, and vice versa. Instead of swiping towards a fixed upper or lower area on the screen, as Sasaki et al. did, participants could be required to freely drag the image along a vertical line which would allow for measurement of the distance from the centre of the screen to the place where the emotional stimulus was dragged to. Then the participants would rate the valence of the stimulus. Based on the current findings it would be hypothesised that the upper/lower the stimulus is located on the vertical axis on the screen, the more positive/negative it would be rated. This finding would support the claim made by Sasaki et al. (2015) that close temporal associations between somatic information and visual events leads to their retrospective integration and provide further credibility to the findings reported herein.

While the emotions *joy* and *sadness* have distinctive sensorimotor correlates, these correlates are very broad in the case of *surprise*. That is, while clapping of hands and head hanging on contracted chest are some of the bodily correlates of *joy* and *sadness* respectively (see Wallbott, 1998), *surprise* manifests in visual search, eye-brow raising, eye-widening, jaw drop, among others (see Reisenzein et al., 2012). However, given that *surprise* seems to be a neutral emotion, its bodily and sensorimotor correlates can be difficult to pinpoint and this situation could lead this emotion to not be regarded as an emotion but as a cognitive state

(Reisenzein et al., 2012). Given current theories arguing that there are degrees in the embodiment of language and emotions (e.g. Chatterjee, 2010; Marmolejo-Ramos & Dunn, 2013; Meteyard, Rodríguez, Bahrami, & Vigliocco, 2012), it is possible that as the more neutral a concept (and the object it refers to) becomes, the lower the degree of sensorimotor properties. Such low activation of sensorimotor correlates and neutral valence can be metaphorically mapped onto space in vertical locations that near the middle instead of upper or lower areas. Moreover, the metaphorical mapping of emotions onto space has so far been limited to the two-dimensional space (i.e. up-down in the *Y* Cartesian coordinate and left-right in the *X* coordinate). It is reasonable to suggest that if valenced concepts were to be allocated in a three-dimensional physical space, highly positively valenced concepts would be placed near the body, highly negatively concepts would be placed far away from the body, and neutrally valenced concepts mid-way between these two. That is, valenced concepts should also have different locations on the *Z* Cartesian coordinate. This is merely conjectural and further empirical testing is needed in order to explore this notion.

Compliance with Ethical Standards

Conflict of Interest: FM-R designed the experiments and analysed the data. All authors discussed the paper, collected data, and wrote the paper. The authors declare no competing interests.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

References

Altarriba, J. & Bauer, L. M. (2004). The distinctiveness of emotion concepts: a comparison between emotion, abstract, and concrete words. *American Journal of Psychology, 117*(3), 389–410

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behaviour Research Methods, Instruments & Computers, 31* (4), 578-602.

Ansorge, U., & Bohner, G. (2013). Investigating the association between valence and elevation with an implicit association task that requires upward and downward responding. *Universitas Psychologica, 12*(5), 1453–1471. doi:10.11144/Javeriana.UPSY12-5.iave

Ansorge, U., Khalid, S., & König, P. (2013). Space-valence priming with subliminal and supraliminal words. *Frontiers in Psychology, 4*:81. DOI: 10.3389/fpsyg.2013.00081

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–45.

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science, 13*(4), 370–374.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B, 57*, 289–300.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15* (11), 527-536.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33* (1), 73-79.

Borghi, A., & Binkofski, F. (2014). *Words as Social Tools: An Embodied View on Abstract Concepts*. New York: Springer

Bradley, M.M., & Lang, P.J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings*. Technical report C-1, Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46* (3), 904-911.

Casasanto, D. (2009). Embodiment of abstract concepts: good and bad in right- and left-handers. *Journal of Experimental Psychology. General, 138*(3), 351–67.

Casasanto, D. (2011). Different Bodies, Different Minds: The Body Specificity of Language and Thought. *Current Directions in Psychological Science, 20*(6), 378–383.

Clark, H. H., Carpenter, P. A., & Just, M. A. (1973). On the meeting of semantics and perception. In W.G. Chase (Ed.), *Visual information processing* (pp. 311-381). New York: Academic Press.

Cumming, G. (2012). *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis.* New York: Routledge.

Cumming, G., & Finch, S. (2005). Inference by eye. Confidence intervals and how to read pictures of data. *American Psychologist, 60* (2), 170-180.

Damjanovic, L., & Santiago, J. (2016). Contrasting vertical and horizontal representations of affect in emotional visual search. *Psychonomic Bulletin & Review, 23* (1), 62-73.

Evans, N. & Levinson, S. (2009). The myth of language universals: language diversity and its importance for cognitive science. *Behavioral and Brain sciences, 32*, 429-492.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558-565

Grühn, D., & Scheibe, S. (2008). Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age? *Behavior Research Methods, 40* (2), 512-521.

Hauk, O., & Pulvermüller, F. (2011). *The lateralization of motor cortex activation to action-words. Frontiers in Human Neuroscience, 5* (149). DOI:10.3389/fnhum.2011.00149

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review, 14* (3), 436-441.

Holstege, G. (1992). The emotional motor system. *European Journal of Morphology, 30* (1), 67-79.

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140* (1), 14–34.

Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are automatic conceptual cores the gold standard of semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science, 39*, 1764-1801.

Leshinskaya, A., & Caramazza, A. (2016). For a cognitive neuroscience of concepts: moving beyond the grounding issue. *Psychonomic Bulletin and Review.* 10.3758/s13423-015-0870-z

Macedo, L., Cardoso, A., Reisenzein, R., Lorini, E., & Castelfranchi, C. (2009). Artificial surprise. In J. Vallverdú & D. Casacuberta (Eds.), *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence* (pp. 267-291). Hershey: Information Science Reference

Marmolejo-Ramos, F., & D'Angiulli, A. (2014). Current research topics in embodied social cognition. *Cognitive Processing, 15*(3), 235–236.

Marmolejo-Ramos, F., & Dunn, J. (2013). On the activation of sensorimotor systems during the processing of emotionally-laden stimuli. *Universitas Psychologica, 12* (5), 1511-1542.

Marmolejo-Ramos, F., Elosúa, M. R., Yamada, Y., Hamm, N. F., & Noguchi, K. (2013). Appraisal of space words and allocation of emotion words in bodily space. *PloS One, 8*(12), e81688. doi:10.1371/journal.pone.0081688

Marmolejo-Ramos, F., Montoro, P. R., Elosúa, M. R., Contreras, M. J., & Jiménez-Jiménez, W. A. (2014). The activation of representative emotional verbal contexts interacts with vertical spatial axis. *Cognitive Processing, 15*, 253–267.

Marocco, D., Cangelosi, A., Fischer, K., & Belpaeme, T. (2010). Grounding Action Words in the Sensorimotor Interaction with the World: Experiments with a Simulated iCub Humanoid Robot. *Frontiers in Neurorobotics, 31* (4). DOI:10.3389/fnbot.2010.00007

Marozzi, M. (2012). A modified Cucconi test for location and scale change alternatives. *Revista Colombiana de Estadística, 35* (3), 369-382.

Marozzi, M. (2014). The multisample Cucconi test. *Statistical Methods and Applications, 23* (2), 209-227.

McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of box plots. *The American Statistician, 32* (1), 12–16.

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: association between affect and vertical position. *Psychological Science, 15* (4), 243-247.

Meteyard, L., Rodríguez, S., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex, 48* (7), 788-804.

Niedenthal, P. M, Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9* (3), 184–211.

Paivio, A.,Yuille, J. C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*, 1-25.

Reali, F., & Arciniegas, C. (2015). Metaphorical conceptualisation of emotion in Spanish. Two studies on the role of framing. *Metaphor and the Social World, 5* (1), 20-41.

Reisenzein, R. (2009). Emotions as metarepresentational states of mind: Naturalizing the belief–desire theory of emotion. *Cognitive Systems Research, 10* (1), 6-20

Reisenzein, R., Meyer, W.-U., & Niepel, M. (2012). Surprise. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior*, 2nd ed. (pp.564-570). New York: Elsevier.

Rueschemeyer, S-A., Lindemann, O., van Rooj, D., van Dam, W., & Bekkering H. (2010). Effects of intentional motor actions on embodied language processing. *Experimental Psychology, 57* (4), 260-266

Sasaki, K., Yamada, Y., & Miura, K. (2015). Post-determined emotion: motor action retrospectively modulates emotional valence of visual images. *Proceedings of the Royal Society B: Biological Sciences, 282:* 20140690

Sasaki, K., Yamada, Y., & Miura, K. (2016). Emotion biases voluntary vertical action only with visible cues. *Acta Psychologica, 163*, 97-106.

Schubert, T. W. (2005). Your highness: vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology, 89*(1), 1–21.

Schwanenflugel, P. J., & Shoben, E.J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 82-102.

Siakaluk, P. D., Pexman, P. M., Sears, C. R., Wilson, K., Locheed, K., & Owen, W. J. (2008). The benefits of sensorimotor knowledge: Body-object interaction facilitates semantic processing. *Cognitive Science, 32* (3), 591-605.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28* (6), 879-896.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–36.

Xie, J., Huang, Y., Wang, R., & Liu, W. (2015). Affective valence facilitates spatial detection on vertical axis: shorter time strengthens effect. *Frontiers in Psychology, 6* (277). doi: 10.3389/fpsyg.2015.00277

Xie, J., Wang, R., & Chang, S. (2014). The mechanisms of valence-space metaphors: ERP evidence for affective word processing. *PLoS ONE, 9* (6): e99479. DOI: 10.1371/journal.pone.0099479.

Xue, J., Marmolejo-Ramos, F., & Pei, X. (2015). The linguistic context effects on the processing of Body-Object Interaction words: An ERP study on second language learners. *Brain Research, 1613*, 37-48

Yohai, V. (1987). High breakdown-point and high efficiency estimates for regression. *Annals of Statistics, 15* (20), 642–656.

Tables

Table 1. Mean concreteness, imageability, context availability and valence ratings of three emotion words as reported in previous studies.

Emotion word	Mean rating			
	Concreteness	Imageability	Context availability	Valence
Joy	2.37	3.7	5.2	8.60
Surprise	3.24	4.2	4.9	7.47
Sadness	1.82	4.0	5.1	1.61

Note. Altarriba and colleagues (Altarriba et al., 1999; Altarriba & Bauer, 2004) and Bradley and Lang (1999), used the words 'surprised' instead of 'surprise' and 'sad' instead of 'sadness'. Brysbaert et al. (2014) provided ratings for 'joy', 'surprise', 'surprised', 'sad' and 'sadness'. The concreteness ratings were performed on a 5-point Likert scale and were reported in Brysbaert et al (2014) (note the concreteness ratings for the words 'joy', 'surprise' and 'sadness' reported by Altarriba and colleagues were 3, 3, and 3.1, respectively, on a 7-point Likert scale). The imageability and context availability ratings were performed on a 7-point Likert scale and were reported in Altarriba et al. (1999). The valence ratings were performed on a 9-point Likert scale and were reported in Bradley and Lang (1999).

Table 2. Demographic and descriptive statistic information of the participants in Study 1 and 2 (MAD = median absolute deviation).

Language	Handedness and Gender				*Total*	Age	
	Right-handed		Left-handed			Range	Median (MAD)
	Male	Female	Male	Female			
+++ Study 1 (rating task)							
English	5	36	1	8	*50*	19-54	20 (1.48)
Hindi	20	23	1	1	*45*	18-26	22 (1.48)
Japanese	48	40	5	2	*95*	18-21	19 (0)
Spanish	22	7	2	0	*31*	18-26	20 (1.48)
Vietnamese	3	34	15	2	*54*	17-27	19 (0)
German	17	24	4	5	*50*	19-37	23 (1.48)
Total	*115*	*164*	*28*	*18*	***325***		
Total (handedness)	*Right handers = 279*		*Left handers = 46*				
Total (gender)	*Males = 143*		*Females = 182*				
Total age range						*17-54*	
Total average age (MAD)							*20 (1.48)*
+++ Study 2 (word allocation task)							
Language	Handedness and Gender				*Total*	Age	
	Right-handed		Left-handed			Range	Median (MAD)
	Male	Female	Male	Female			
English	10	38	1	2	*51*	19-48	20 (1.48)
Hindi	22	24	1	1	*48*	18-26	22 (1.48)
Japanese	82	33	5	3	*123*	18-23	19 (1.48)
Spanish	11	18	2	2	*33*	18-60	24 (7.41)
Vietnamese	4	37	14	2	*57*	17-27	19 (0)
German	10	28	5	7	*50*	18-45	24.5 (4.44)
Total	*139*	*178*	*28*	*17*	***362***		
Total (handedness)	*Right handers = 317*		*Left handers = 45*				
Total (gender)	*Males = 167*		*Females = 195*				
Total age range						*17-60*	
Total average age (MAD)							*20 (1.48)*

Note. The data were obtained in the following institutions: Teesside University (United Kingdom), G.H. Raisoni College of Engineering (India), Kyushu University (Japan), Universidad Simón Bolívar (Venezuela), Hanoi University (Vietnam), and Leibniz Knowledge Media Research Center (Germany).

Figures

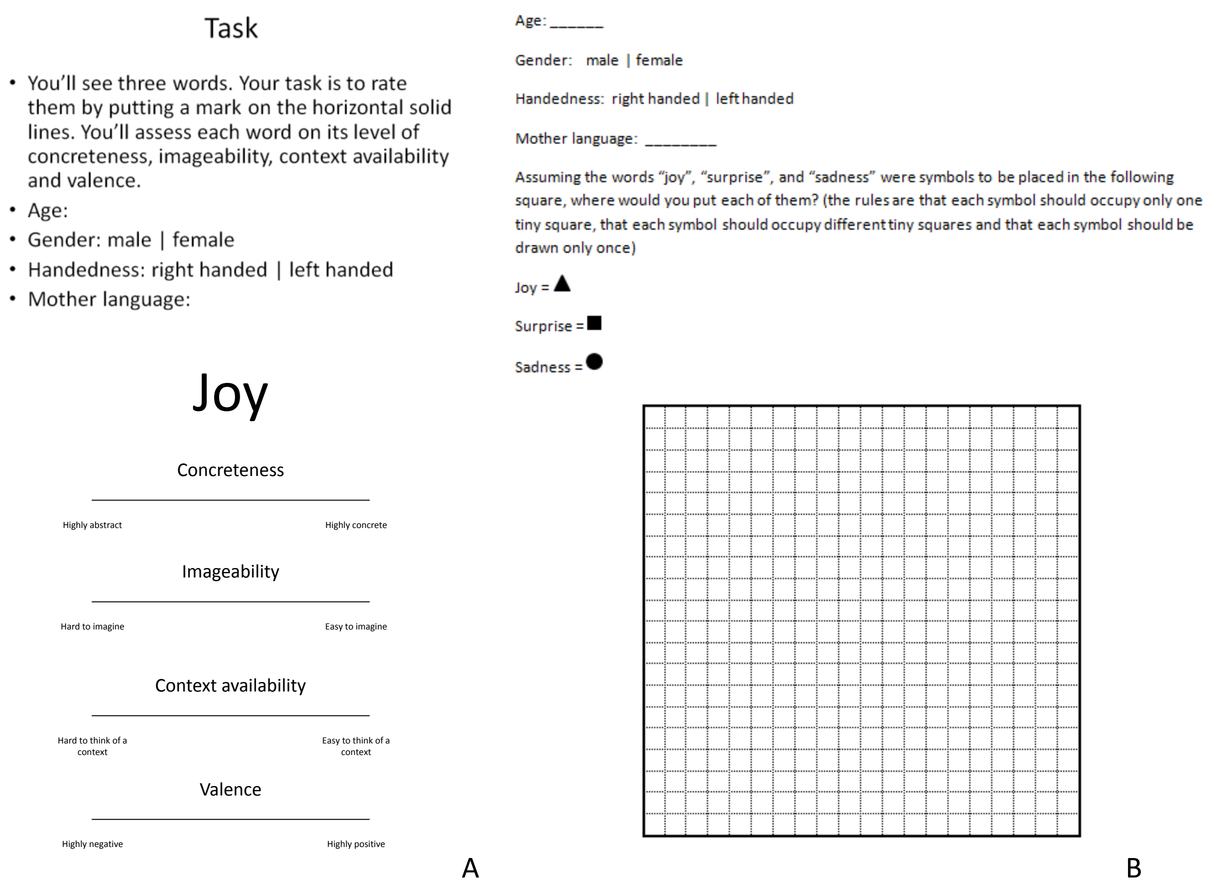

Figure 1. Materials used in the rating (A) and the word allocation (B) tasks. Figure 1A shows the case of *joy* for illustrative purposes only.

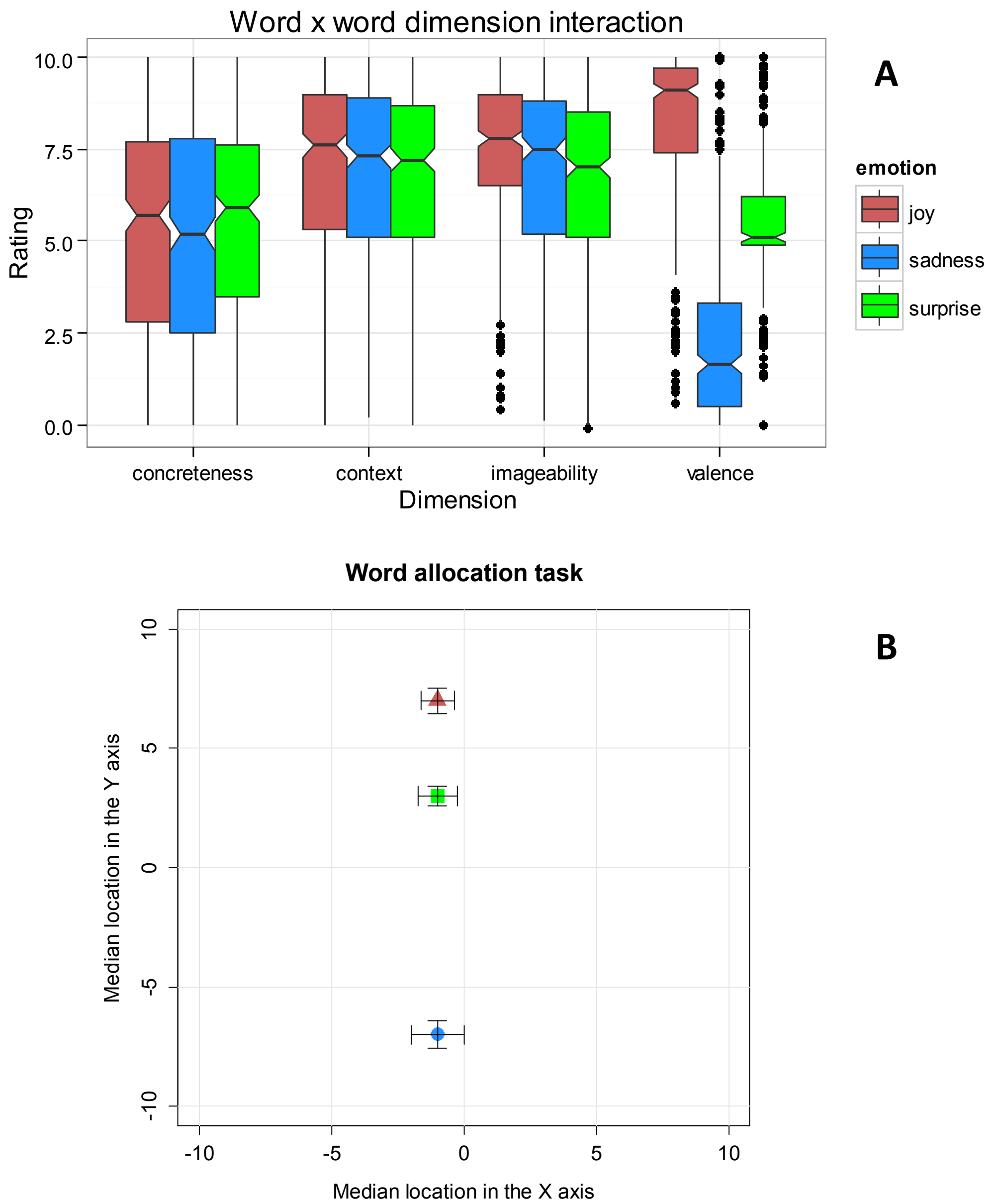

Figure 2. Results of the rating (A) and the word allocation (B) tasks. The notches in the boxplots and the error bars represent 95% CI around the median. ▲=joy, ■=surprise and ●=sadness.

Appendix

- **Supplementary graphical results of the non-significant effects of the factors language and handedness in the word allocation task**

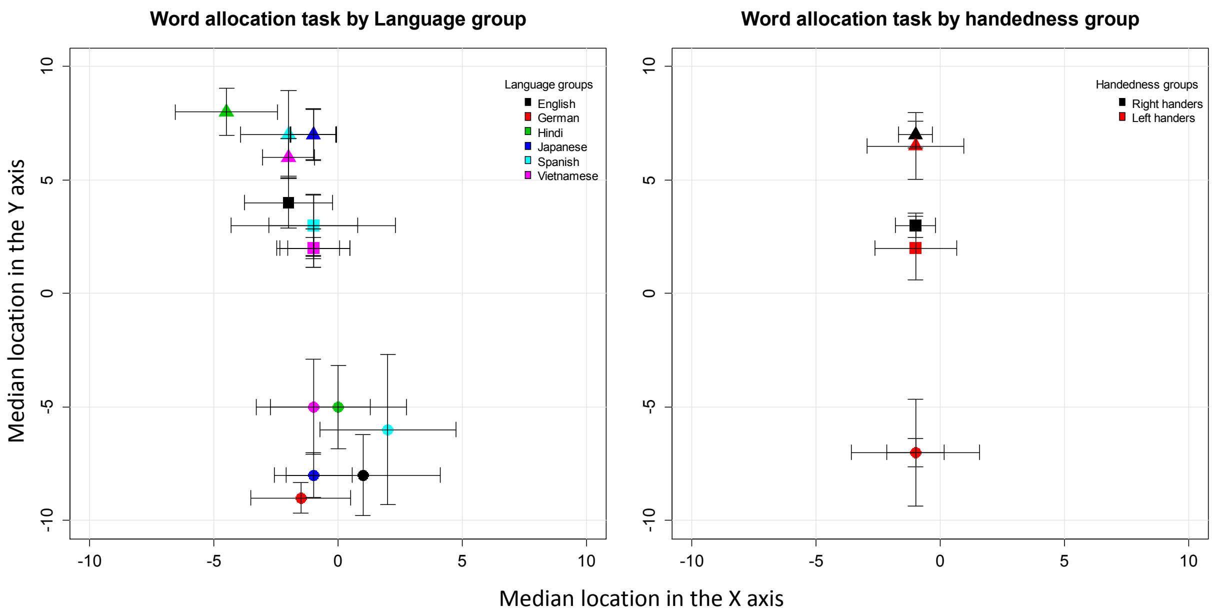

Figure A1. Results of the word allocation task per language and handedness group. The error bars represent 95% CI around the median. ▲=joy, ■=surprise and ●=sadness

- **Supplementary word allocation task data**

Note that in the allocation task reported above, both word order and symbol order were fixed (see Figure 1). That is, the word order was always joy, surprise and sadness and they were paired with a triangle, a square and a circle, respectively. Thus, a follow-up study in which word order (i.e. six possible combinations), symbol order (i.e. also six possible combinations) and their pairings were fully counterbalanced was conducted (i.e. 36 different word order and symbol order combinations, which gave rise to 36 different paper-based word allocation questionnaires).

A total of 473 participants were randomly allocated to each of the 36 questionnaires (see Table A1). Word order and symbol order were added to the same modelling approach used for the analyses of the data from Study 2. The factors were hierarchically entered in this order: participant (P), language (L), gender (G), handedness (H), age (A), word order (Wo), symbol order (So) and word (W).

The results showed that, as found in Study 2, no factor had a significant effect on the X axis: P model: F_r = .16, p=.67, P + L model: F_r = .78, p=.66; P + L + G model: F_r = 1.75, p=.17; P + L + G + H model: F_r = 1.32, p=.24; P + L + G + H + A model: F_r = .06, p=.79; P + L + G + H + A + Wo model: F_r = .27, p=.99; P + L + G + H + A + Wo + So model: F_r = .13, p=.99; and P + L + G + H + A + Wo + So + W model: F_r = 5.07, p=.07. Also, the median X location for the three words was -1: Mdn_{joy} = -1 (95% CI [-1.36, -0.63]), $Mdn_{surprise}$ = -1 (95% CI [-1.50, -0.49]), and $Mdn_{sadness}$ = -1 (95% CI [-1.79, -0.20]).

The analyses also replicated the results in the Y axis shown in Study 2 such that only the model including the factor 'word' was significant: P model: F_r = .10, p=.75, P + L model: F_r = .57, p=.74; P + L + G model: F_r = 1.62, p=.19; P + L + G + H model: F_r = .01, p=.92; P + L + G + H + A model: F_r = 1.27, p=.25; P + L + G + H + A + Wo model: F_r = .37, p=.99; P + L + G + H + A + Wo + So model: F_r = .86, p=.97; and P + L + G + H + A + Wo + So + W model: F_r = 574.37, p<.001. The median locations for the three words differed: Mdn_{joy} = 7 (95% CI [6.49, 7.50]), $Mdn_{surprise}$ = 3 (95% CI [2.56, 3.43]), and $Mdn_{sadness}$ = -7 (95% CI [-7.72, -6.27]).

Analyses of the effects of the covariates P, L, G, H, A, Wo, and So on the X data for each of the three words showed an effect of L in the allocation of the word *joy* (P + L model: F_r = 7.58, p=.01) such that some languages placed this word more rightward/leftward than others (all other models in this word and the words *surprise* and *sadness* had p>.05).

Analyses of the effects of the same covariates on the Y data for each of the three words showed effects of P, L and A in the allocation of the word *sadness* (P model: $F_r = 8.97$, $p=.002$; P + L model: $F_r = 18.76$, $p=5.86e^{-05}$; and P + L + G + H + A model: $F_r = 7.69$, $p=.004$) such that some participants, languages and age groups allocated this word more upward/downward than others (all other models in this word and the words *surprise* and *joy* had $p>.05$).

+++ Study 3 (word allocation task)

Language	Handedness and Gender				*Total*	Age	
	Right-handed		Left-handed			Range	Median (MAD)
	Male	Female	Male	Female			
Japanese	43	17	4	1	*65*	18-35	19 (1.48)
Spanish	93	226	3	16	*338*	16-57	21 (4.44)
German	18	43	4	5	*70*	19-48	25 (2.96)
Total	154	286	11	22	***473***		
Total (handedness)	*Right handers = 440*		*Left handers = 33*				
Total (gender)	*Males = 165*		*Females = 308*				
Total age range						*16-57*	
Total average age (MAD)							*21 (4.44)*

Table A1. Demographic and descriptive statistic information of the participants in a follow-up study of the word allocation task (MAD = median absolute deviation).