COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 43 (2019) e12727

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12727

Is Emotional Magnitude Spatialized? A Further Investigation

Kevin J. Holmes,[a] Candelaria Alcat,[a] Stella F. Lourenco[b]

[a]*Department of Psychology, Colorado College*
[b]*Department of Psychology, Emory University*

Received 20 January 2018; received in revised form 4 March 2019; accepted 11 March 2019

Abstract

Accumulating evidence suggests that different magnitudes (e.g., number, size, and duration) are spatialized in the mind according to a common left–right metric, consistent with a generalized system for representing magnitude. A previous study conducted by two of us (Holmes & Lourenco, 2011) provided evidence that this metric extends to the processing of *emotional magnitude*, or the intensity of emotion expressed in faces. Recently, however, Pitt and Casasanto (2018) showed that the earlier effects may have been driven by a left–right mapping of mouth size rather than emotional magnitude, and they found no evidence for an emotional magnitude mapping when using words as stimuli. Here, we report two new experiments that further examine these conclusions. In Experiment 1, using face stimuli with mouths occluded, we replicate the original finding: Less emotional faces were associated with the left and more emotional faces with the right. However, we also find that people can reliably infer the sizes of the occluded mouths, and that these inferred mouth sizes can explain the observed left–right mapping. In Experiment 2, we show that comparative judgments of emotional words yield a left–right mapping of emotional magnitude not attributable to stimulus confounds. Based on these findings, we concur with Pitt and Casasanto that faces pose challenges for isolating the forces driving spatialization, but we suggest that emotional magnitude, when assessed using unconfounded stimuli in a sufficiently sensitive task, may indeed be spatialized as originally proposed. Suggestions for further research on the spatialization of emotional magnitude are discussed.

Keywords: Magnitude; Emotion; Space; Intensity; Valence; Faces; Words; General magnitude system

1. Introduction

Magnitude information is ubiquitous in human experience, whether embedded in numerical symbols such as Arabic digits or in non-numerical cues such as object size

Correspondence should be sent to Kevin J. Holmes, Department of Psychology, Colorado College, 14 E. Cache La Poudre St., Colorado Springs, CO 80903. E-mail: kjholmes@coloradocollege.edu

or event duration. Some investigators have proposed that different magnitudes (e.g., number, size, and duration) are processed as part of an integrated system that draws on common neural and cognitive resources (Bonn & Cantlon, 2017; Bueti & Walsh, 2009; Lourenco, 2016; Walsh, 2003). Support for such a *general magnitude system* comes from evidence of overlapping activation in parietal cortex (Cantlon, Platt, & Brannon, 2009; Cohen Kadosh, Lammertyn, & Izard, 2008) and of cognitive interactions between different magnitudes (e.g., Henik & Tzelgov, 1982; Lourenco, Ayzenberg, & Lyu, 2016).

Additional support for a general magnitude system comes from studies showing that people implicitly associate smaller and larger magnitudes with the left and right sides of space, respectively. This left-to-right orientation of increasing quantities—initially regarded as specific to number (i.e., the classic Spatial-Numerical Association of Response Codes [SNARC] effect; Dehaene, Bossini, & Giraux, 1993; for a meta-analysis, see Wood, Willmes, Nuerk, & Fischer, 2008)—has also been observed for non-numerical magnitudes, including size (Ren, Nicholls, Ma, & Chen, 2011; Sellaro, Treccani, Job, & Cubelli, 2014), duration (Fabbri, Cancellieri, & Natale, 2012; Vallesi, Binns, & Shallice, 2008), and even abstract dimensions such as risk (Macnamara, Loetscher, & Keage, 2019), and may be supported by shared topographic organization in parietal cortex (Harvey, Fracasso, Petridou, & Dumoulin, 2015). Such findings suggest that different magnitudes may be spatialized according to a common mental metric (see also Holmes & Lourenco, 2013).

A previous series of experiments conducted by two of us (Holmes & Lourenco, 2011; henceforth H&L) probed whether this common metric extends to the processing of emotion—specifically, the intensity of emotion expressed in faces (which H&L termed "emotional magnitude"). When participants responded to photographs of faces that varied in emotional magnitude, their reaction times (RTs) revealed that less emotional faces were implicitly associated with the left side of space and more emotional faces with the right. This spatialization of emotional magnitude, which mirrors the left-to-right orientation of number and other magnitudes (e.g., Dehaene et al., 1993; Ren et al., 2011), was observed regardless of whether the valence of the emotion was positive (happy) or negative (angry). H&L proposed that these findings might reflect the workings of a "hyper-general" system of magnitude representation—one that supports the processing of a range of prothetic cues characterized by analog variation (i.e., more vs. less; Stevens, 1957), even emotional stimuli.

Recently, Pitt and Casasanto (2018; henceforth P&C) provided an alternative explanation for H&L's findings. They pointed out that H&L's face stimuli varied not only in emotional magnitude, but also in the size of a salient feature: the mouth. Indeed, emotional magnitude and mouth size were highly correlated in H&L's stimuli; the more emotional faces had, on average, larger mouths. In reanalyses of H&L's data controlling for mouth size, P&C found that emotional magnitude no longer significantly predicted participants' RTs. Notably, however, neither did mouth size when controlling for emotional magnitude. Although these results suggest that neither variable was solely responsible for H&L's effects, P&C argued that mouth size was the more likely candidate for spatialization. This argument was based on other research showing a left–right mapping of object size (e.g., Ren et al., 2011), not on a direct test disentangling the effects of the two variables. P&C suggested that such a test may be impossible because mouth size is perceptually integral to (i.e., inseparable from) emotional magnitude.

Although the effects of emotional magnitude and mouth size may indeed be inseparable in normal face perception (and in H&L's data), they could, in principle, be distinguished experimentally. In the present research, we attempted to isolate the effect of emotional magnitude by occluding the mouths in H&L's face stimuli. In Experiment 1 (an otherwise direct replication of one of H&L's experiments), participants classified the gender of mouth-occluded faces whose expressions varied in emotional magnitude as well as in emotional valence, for which a left–right mapping in right-handers has been documented (see P&C; Casasanto, 2014). If emotional magnitude is spatialized independent of mouth size (and of the documented valence mapping), participants should associate less emotionally intense faces with the left and more emotionally intense faces with the right, controlling for the emotional valence of the faces. Even with the mouth occluded, however, other parts of the face (e.g., cheeks, chin, etc.) might license inferences about mouth size. Therefore, for any spatial mapping of face stimuli to be attributed to emotional magnitude per se, the mapping must be shown to be independent of *inferred*, not just actual, mouth size.

The challenges presented by the use of face stimuli, some of which we have just outlined, led P&C to rely on words instead. In addition to their reanalyses of H&L's data, P&C reported two new experiments in which right-handed participants used left and right response keys to classify various words (e.g., *gregarious*, *obstinate*) as expressing a positive or negative emotion (valence task) or as expressing a more intense or less intense emotion (intensity task). However, the valence task yielded RTs consistent with the documented left-right emotional valence mapping, the intensity (magnitude) task yielded no evidence of a left-right emotional magnitude mapping. We suspected, however, that P&C's task and stimuli may not have been sufficiently sensitive to reveal the latter mapping. In Experiment 2, we assessed this mapping using a magnitude comparison task in which participants judged the intensity of each word relative to a comparison word—rather than leaving the reference point unspecified, as in P&C's experiments—to ensure that the judgment was unambiguous. Our word stimuli were also selected to highlight contrasting levels of emotional intensity, arguably rendering intensity more salient than in P&C's stimuli. To preview the results, we find that participants associate less emotionally intense faces and words with the left side of space and more emotionally intense faces and words with the right, but that the effect for faces can be explained by inferred mouth size. The effect for words, which lack the stimulus confounds inherent in faces, is consistent with H&L's proposed emotional magnitude mapping.

2. Experiment 1: Mouth-occluded faces

2.1. Method

2.1.1. Participants

Forty-eight undergraduate students, a sample size comparable to that of H&L's Experiment 2 ($N = 46$), participated for payment. All participants had normal or corrected-to-normal vision and gave written informed consent. The majority (46) were right-handed ($M = 73.0$, ranging from −38.5 to 100 according to the Edinburgh Handedness Inventory, EHI; Oldfield, 1971). Three additional participants were excluded for mean accuracy ($n = 1$) or mean RT ($n = 2$) >2.5 standard deviations (*SD*s) from the group means.

2.1.2. Materials

The face stimuli (90 × 65 mm, 10.3° × 7.4°) were identical to those of H&L's Experiment 2, except that a rectangle was centered over the mouth area of each face, including lips (see Fig. 1). The rectangle's size (30 × 10.8 mm, 3.4° × 1.2°) was the same across faces and was selected to fully occlude the mouth area of the faces with the largest mouths. Stimuli were obtained from the NimStim set (Tottenham et al., 2009) and converted into grayscale. There were 30 unique face images, with six actors (three male, three female) exhibiting five distinct emotional expressions (which we labeled *neutral*, *happy*, *angry*, *extremely happy*, and *extremely angry*).

For these mouth-occluded faces, we separately collected ratings of emotional magnitude (EM), emotional valence (EV), and mouth size from participants on Amazon Mechanical Turk. These participants rated either (a) EM ($n = 50$), (b) EV ($n = 52$), or (c) estimated mouth size ($n = 50$) for all 30 faces (presented in random order) on a 7-point scale (EM: 1 = neutral; 7 = very emotional; EV: 1 = very negative; 7 = very positive; estimated mouth size: 1 = mouth fully closed, lips touching; 7 = mouth fully open, lips apart). Table 1 shows the mean ratings for each expression.

For analyses of the ratings, we used the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2018). EM, EV, and estimated mouth size ratings all

Fig. 1. Mouth-occluded face stimuli from Experiment 1, shown here for one of the six models.

Table 1
Mean (*SD*) ratings of face stimuli in Experiment 1

	Emotional Expression				
	Neutral	Happy	Angry	Extremely Happy	Extremely Angry
Emotional magnitude (EM)	2.38 (0.21)	4.58 (0.28)	4.76 (0.38)	5.59 (0.33)	5.79 (0.50)
Emotional valence (EV)	3.79 (0.08)	5.87 (0.21)	2.28 0.17)	6.27 (0.15)	1.71 (0.36)
Estimated mouth size	2.67 (0.23)	4.73 (0.32)	2.97 (0.43)	5.94 (0.31)	4.03 (0.93)

varied significantly across expressions (with random intercepts for actors; EM: $\chi^2(4) = 83.66$; EV: $\chi^2(4) = 134.77$; estimated mouth size: $\chi^2(4) = 64.26$; all *p*s < .0001). Mouth size estimates predicted both EM and EV ratings (with random intercepts for actors; EM and estimated mouth size: $r = .65$, $\chi^2(1) = 16.48$, $p < .0001$; EV and estimated mouth size: $r = .64$, $\chi^2(1) = 15.98$, $p < .0001$), but neither EM nor EV ratings predicted the other ($r = -.04$, $\chi^2(1) = .05$, $p = .83$).

2.1.3. Procedure

The procedure was identical to that of H&L's Experiment 2. Participants classified the gender of each face by pressing left ("Q") or right ("P") computer keys with the corresponding index finger. There were two blocks of trials, with order counterbalanced: one in which participants pressed the left key for male and the right key for female, and the other with the reverse key assignments. Each block consisted of 10 practice trials and 90 test trials. Across the test trials of each block, each of the 30 faces was presented three times (and each of the five expressions 18 times) in random order. On each trial, a black fixation cross was presented centrally on a white background for 500 ms, followed by a face that remained onscreen until participants made a response. The next trial began 500 ms after the response. Instructions emphasized speed and accuracy. Procedures were approved by the local ethics board.

2.2. Results

Trials in which participants responded incorrectly (3.5%) or in which reaction time (RT) was >2.5 standard deviations (*SD*s) from participant means (2.7%) were excluded. On the remaining trials, mean RT was 531 ms ($SD = 67$).

2.2.1. Emotional magnitude

To assess the spatialization of emotional magnitude, we entered RTs into a mixed-effects model with response side and EM ratings as predictors, and with random slopes and intercepts for participants and actors. In this model, the interaction between response side and EM ratings served as an index of left-to-right orientation of emotional magnitude. This interaction was significant, $\chi^2(1) = 5.07$, $p = .02$. Fig. 2A displays the mean difference in RT between right-side and left-side responses (dRT: right – left) for each emotional expression. As shown in the figure, less emotional faces were associated with the left side of space and more emotional faces with the right, replicating H&L. (The effect of emotional magnitude was also significant when assessed using participant-specific regression slopes [see Supplementary Material]—the approach used by H&L, following research on the spatialization of numerical magnitude [Fias & Fischer, 2005]; see P&C for discussion).

2.2.2. *Emotional valence*

To assess whether the data from our predominantly right-handed sample also reflected the previously documented left–right mapping of emotional valence in right-handers (see P&C), we conducted an analysis analogous to that for emotional magnitude, using EV ratings rather than EM ratings as a predictor. The interaction between response side and EV ratings—an index of the emotional valence mapping—trended toward significance, $\chi^2(1) = 2.96$, $p = .085$, with more negative and more positive faces tending to be associated with left and right, respectively (see Fig. 2B). To evaluate whether the contributions of emotional magnitude and emotional valence to participants' RTs were independent of each other, we residualized EM ratings with respect to EV ratings, and vice versa, and

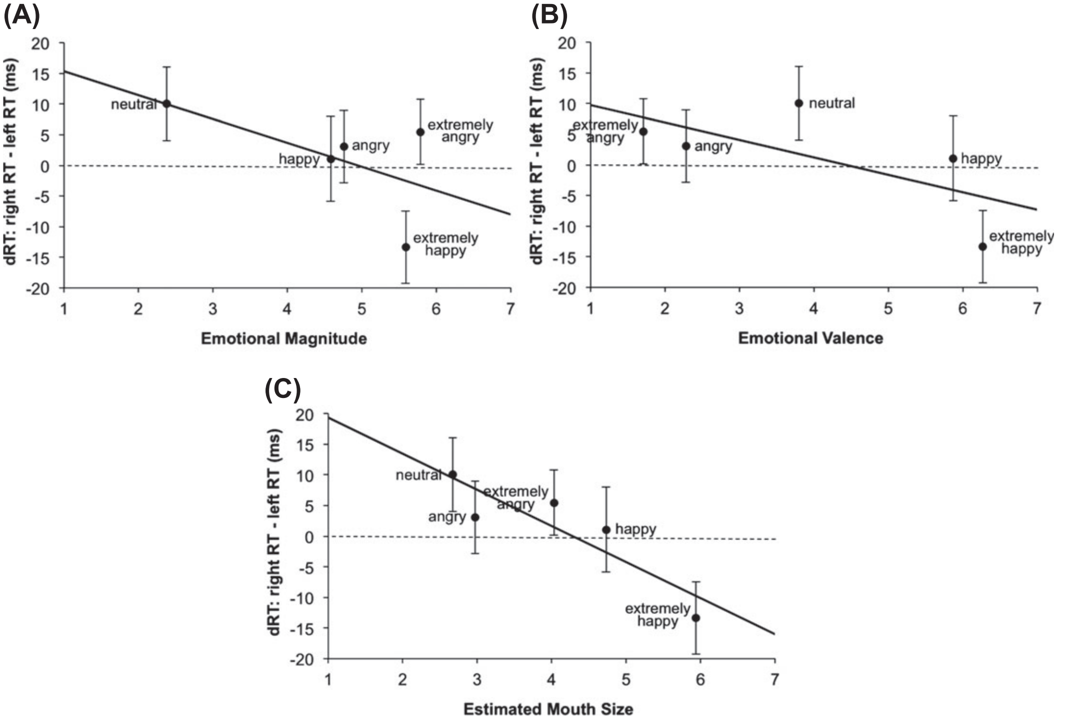

Fig. 2. Mean dRT plotted by (A) mean EM ratings, (B) mean EV ratings, and (C) mean estimated mouth sizes for each expression. Negative dRT values indicate faster right-side responses, and positive dRT values indicate faster left-side responses. Error bars represent *SEM*.

we conducted separate mixed-effects analyses with each set of residuals as a predictor. The effect of emotional magnitude remained significant when controlling for emotional valence, $\chi^2(1) = 4.65$, $p = .03$, and the effect of emotional valence remained marginal when controlling for emotional magnitude, $\chi^2(1) = 3.15$, $p = .076$. (When assessed using participant-specific regression slopes, the effect of emotional valence was significant both with and without controlling for emotional magnitude; see Supplementary Material.)

2.2.3. *Inferred mouth size*

Given that mouth size estimates covaried with both EM and EV ratings, it is possible that participants inferred the sizes of the occluded mouths during the task, and that these inferred mouth sizes can explain the effects of emotional magnitude and emotional valence reported above. To evaluate this possibility, we conducted an analysis analogous to those reported above, using mouth size estimates (instead of EM or EV ratings) to predict RTs. The interaction between response side and mouth size estimates was significant, $\chi^2(1) = 4.37$, $p = .04$, indicating that faces with smaller estimated mouth sizes were associated with the left and faces with larger estimated mouth sizes were associated with the right (see Fig. 2C). To evaluate whether the effects of emotional magnitude and emotional valence were independent of this inferred mouth size mapping, we residualized EM and EV ratings separately with respect to mouth size estimates, and we conducted mixed-effects analyses with each set of residuals as a predictor. Neither the effect of emotional magnitude nor the effect of emotional valence was significant when controlling for mouth size estimates (EM: $\chi^2(1) = 0.24$, $p = .63$; EV: $\chi^2(1) = 0.04$, $p = .83$). However, when assessed in an analogous fashion, the effect of mouth size estimates trended toward significance when controlling for emotional magnitude, $\chi^2(1) = 3.01$, $p = .082$, and when controlling for emotional valence, $\chi^2(1) = 3.47$, $p = .062$. (Similar patterns were observed when using participant-specific regression slopes; see Supplementary Material.)

2.3. *Discussion*

When judging the gender of mouth-occluded faces, participants associated less emotional faces with the left side of space and more emotional faces with the right, consistent with the emotional magnitude mapping proposed by H&L. However, left-side responses were also faster for faces with more negative expressions and with smaller estimated mouth sizes, and right-side responses were faster for faces with more positive expressions and with larger estimated mouth sizes. The effect of emotional magnitude remained significant when controlling for emotional valence, and the effect of emotional valence remained significant when controlling for emotional magnitude, but neither effect held when controlling for estimated mouth size.

So what exactly was spatialized? We suggest two possibilities. First, even though mouth size estimates (from a separate sample) can explain the observed effects of emotional magnitude and valence, it is unclear whether participants actually inferred the sizes of the occluded mouths while completing the task. However, given that other participants were able to make mouth size estimates that varied reliably across expressions (likely by relying on clues from the visible parts of the face), it appears that mouth size was readily inferable from our stimuli, so participants may very well have made such inferences during the task. Indeed, because controlling for estimated mouth size eliminated the effects of emotional magnitude and valence, the results can be fully explained by spatialization of the mouth sizes that participants putatively inferred. This possibility is in line with P&C's reanalysis suggesting that (actual) mouth size was spatialized in H&L's experiments, and with other studies showing spatialization of size information (e.g., Ren et al., 2011).

The second possibility is that emotional magnitude and valence are represented—and spatialized—*via* mouth size. That is, the size of the mouth (actual or inferred) may communicate critical information about emotion—so much so that controlling for its contribution eliminates the spatialization of key emotional properties such as magnitude and valence. According to this possibility, attempts to deconfound the contributions of emotional magnitude and valence from that of mouth size (or other spatial properties of the face) will inevitably be futile, since the former are constituted by the latter.[1] Thus, despite our efforts to isolate emotional magnitude from other factors, the results of Experiment 1 do not afford clear conclusions regarding the forces driving spatialization. Although studies using stimuli other than faces have provided more definitive evidence for the spatialization of emotional valence (see P&C) and of size (e.g., Ren et al., 2011), whether emotional magnitude is spatialized remains unclear.

In Experiment 2, we took a different approach to investigating this question—the same one used by P&C to avoid the confounds inherent in faces. As mentioned above, when using words as stimuli, P&C found no evidence for a left–right mapping of emotional magnitude (which they referred to as "intensity"). However, two aspects of P&C's experiments may have obscured such a mapping. First, P&C's participants judged whether each word expressed a "more intense emotion" or "less intense emotion," but without a clear standard for comparison (i.e., more or less intense than *what*?). P&C's valence judgments, in contrast, were unambiguous; when judging whether a word expresses a positive or negative emotion, there is a stable, implied reference point (neutral valence). We hypothesized that judging intensity relative to an explicit reference point (i.e., more or less intense than another word) might yield a left–right mapping of intensity—as suggested by studies of the spatialization of numerical magnitude, in which judgments made relative to a reference number (e.g., more or less than 5) often yield stronger spatial mappings than other kinds of judgments (Wood et al., 2008). Second, several of P&C's words convey relatively subtle information about emotional intensity. Although the intensity of words like *gregarious*, *obstinate*, and *attractive* can be discerned when task-relevant (indeed, P&C provided evidence that participants computed the relative intensities of their stimuli), these words carry other aspects of meaning that are arguably more salient. We suspected that words chosen specifically to capture contrasting levels of intensity (e.g., *excited* and *calm*), as in studies of the conceptual structure of emotion (e.g., Russell, 1980), might lead participants not only to compute intensity, but also to spatialize it.

We investigated these possibilities in Experiment 2 by adapting P&C's word intensity judgment task. Participants judged intensity relative to a reference word for a set of emotional words that differed noticeably in intensity, as well as in valence. If intensity (i.e., emotional magnitude) is spatialized in this context, then participants should associate less intense words with the left and more intense words with the right, and this effect should be independent of any spatialization of the words' valence.

3. Experiment 2: Comparative judgments of intensity-salient words

3.1. Method

3.1.1. Participants

Forty-eight undergraduate students participated for payment. We opted for a larger sample size than in P&C's experiments (n = 16 in Experiment 1a and 32 in Experiment 1b, for their intensity task) to ensure sufficient power and for comparison with Experiment 1. All participants had normal or corrected-to-normal vision and gave written informed consent. As in P&C's experiments, all participants were right-handed (EHI: M = 90.2; range: 70–100). One additional participant was excluded for a mean RT >2.5 *SD*s from the group mean.

3.1.2. Materials

Stimuli were selected from among the 28 emotional adjectives used in studies of the conceptual structure of emotion by Russell and colleagues (e.g., Russell, 1980; Russell, Weiss, & Mendelsohn, 1989). Based on a norming task with a separate sample of Amazon Mechanical Turk participants (N = 40), modeled after that of P&C's Experiment 1a, we selected eight words whose mean standardized intensity ratings differed significantly from that of a neutral comparison word (*glad*, .05): four rated as more intense (*tense*, .34; *happy*, .48; *alarmed*, .64; *excited*, .81), and four rated as less intense (*bored*, −.95; *calm*, −.86; *tired*, −.77; *relaxed*, −.60; all ps < .05). Within both the more intense and less intense sets, two words were positive and two were negative, according to valence ratings completed as part of the same norming task (see also Russell, 1980). The mean standardized valence ratings of the words were as follows: *alarmed*, −.41; *bored*, −.36; *calm*, 0.71; *excited*, 1.11; *happy*, 1.17; *relaxed*, 0.78; *tense*, −0.47; *tired*, −0.33. Thus, for the eight stimuli, intensity was fully crossed with valence.

3.1.3. Procedure

On each trial, one of the eight words appeared centrally in black text on a gray background. Participants judged whether the word expressed a "more intense emotion" or "less intense emotion" relative to the word *glad*, by pressing left ("Q") or right ("P") computer keys with the corresponding index finger. The next trial began 500 ms after the response. There were two blocks of trials, with order counterbalanced: one in which participants pressed the left key for more intense and the right key for less intense, and the other with the reverse key assignments. Each block consisted of 16 practice trials (2 per word) and 128 test trials (16 per word), presented in random order. Instructions emphasized speed and accuracy. Following this RT task, participants were asked to rate (a) the intensity of the eight words (from 0, least intense, to 10, most intense), and (b) the valence of the words (from −5, most negative, to 5, most positive), presented individually in random order. Intensity and valence ratings were completed in separate blocks, with order counterbalanced. Procedures were approved by the local ethics board.

3.2. Results and discussion

The eight words varied significantly in their intensity ratings, $F(7, 329) = 133.65$, $p < .0001$, and their valence ratings, $F(7, 329) = 348.04$, $p < .0001$. Mean intensity and mean valence ratings were not correlated across words, $r = -.08$, $p > .8$.

For the RT data, trials with RTs >2.5 *SD*s from participant means (4.0%) were excluded. On the remaining trials, mean RT was 731 ms ($SD = 154$). Following P&C, to assess the spatialization of intensity, we entered RTs into a mixed-effects model with response side and participant-specific intensity ratings as predictors, and with random slopes and intercepts for participants. The interaction between response side and intensity ratings was significant, $\chi^2(1) = 4.20$, $p = .04$. Fig. 3 displays the mean dRT (right – left) for each word, showing that participants associated less intense words with the left side of space and more intense words with the right. This mapping was not driven by the valence of the words; as shown in the figure, the two most negatively valenced words in the set (*tense* and *alarmed*) were the ones most strongly associated with the right, contrary to right-handers' tendency to associate negative stimuli with the left in other tasks (see P&C).

To statistically control for any effect of valence, we residualized intensity ratings with respect to valence ratings for each participant and entered the residuals as a predictor in a mixed-effects model, as in Experiment 1. The interaction between response side and residualized intensity ratings was significant, $\chi^2(1) = 4.77$, $p = .03$, indicating that the effect of intensity was independent of valence. (The intensity effect was also significant, both with and without controlling for valence, when assessed using participant-specific

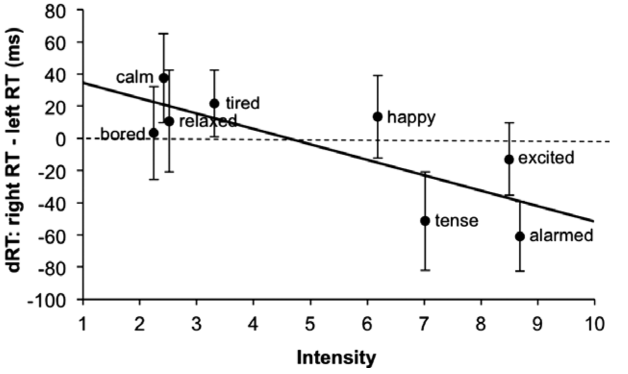

Fig. 3. Mean dRT plotted by mean intensity ratings for each word. Error bars represent *SEM*.

regression slopes. The valence effect was nonsignificant when assessed directly, both with and without controlling for intensity; see Supplementary Material.)

These results show that participants associated less intense words with the left and more intense words with the right, independent of any spatialization of valence. This finding stands in contrast to P&C's results, which showed no intensity mapping for words. Unlike P&C's word intensity task, ours required participants to judge intensity relative to a reference word and included a set of words for which intensity was arguably more salient. Either or both of these distinguishing features may account for the different outcomes. Notably, P&C's experiments were motivated by the possibility that mappings of emotional magnitude/intensity and valence are both stored in long-term memory, but that "at any moment only one of these contradictory mappings is activated strongly enough to produce the predicted pattern of behavior" (p. 4). Our findings suggest that intensity judgments alone (as in P&C's experiments) are not sufficient for spatialization, and that an explicit reference point and/or intensity-salient stimuli may be needed. In contrast, the emotional valence mapping has been demonstrated using a variety of tasks and stimuli (Casasanto, 2014), perhaps in part because the valence of individual stimuli can be discerned unambiguously due to the implied, neutral-valence reference point. For this reason, the valence mapping may be less context-dependent than the intensity mapping (see below for further discussion of the extent to which our results support the proposed intensity mapping).

4. General discussion

Our experiments extend the work of H&L and P&C by providing new data that advance the debate on the spatialization of emotional magnitude. Using mouth-occluded face stimuli, we replicated H&L's finding that less emotional faces were associated with the left side of space and more emotional faces with the right, independent of the valence of the expressions. However, we also found that emotional magnitude covaried with estimates of the sizes of the occluded mouths, and that these estimates—which reflect inferences participants may have made when judging the faces—account for the observed left–right mapping. In light of these results, we concur with P&C that faces pose insuperable challenges for isolating the forces driving spatialization.

Turning to word stimuli that avoided these stimulus confounds, we found that less emotionally intense words were associated with the left and more emotionally intense words with the right, independent of valence—consistent with the proposed emotional magnitude (intensity) mapping. These results suggest that emotional magnitude, when assessed using unconfounded stimuli in a sufficiently sensitive task, may indeed be spatialized as originally proposed by H&L. Admittedly, however, the results do not provide unambiguous support for spatialization. As with most binary classification tasks, the findings are also consistent with an alternative explanation based on *polarity correspondence* (Proctor & Xiong, 2015). Participants responded faster when the poles of one of the stimulus dimensions (intensity: more/less) and the response dimension (left/right) corresponded (i.e., +polar: *more intense* and *right*; -polar: *less intense* and *left*) than when they did not (i.e., *more intense* and *left*; *less intense* and *right*)—a congruity effect that may reflect structural overlap between the dimensions, rather than spatialization per se. This explanation can account for P&C's emotional valence effect for words, as well as for much of the evidence for spatialization of number and other magnitudes (e.g., Dehaene et al., 1993; Ren et al., 2011), given that these studies also relied on binary classification.

Only recently have manipulations been devised to disambiguate between spatialization and polarity correspondence accounts of RT congruity effects in binary classification tasks (e.g., Santiago & Lakens, 2015; Song, Chen, & Proctor, 2017). Future research could adapt these manipulations to the dimension of emotional magnitude. One possible approach would be to replicate Experiment 2 while varying keyboard eccentricity (i.e., the lateral placement of the keyboard), a factor that has been shown to modulate classic RT congruity effects, presumably by changing the polarity of the response dimension (e.g., *left* = +polar when the keyboard is on the left; Proctor & Xiong, 2015). If a left–right mapping of emotional magnitude is observed regardless of eccentricity, as has been found for number (Santiago & Lakens, 2015), this would argue against polarity correspondence as an explanation for the mapping.

Based on their findings, P&C concluded that "if people spatialize emotional intensity at all, it is unlikely that they activate a left–right mapping of intensity with the same strength or automaticity" (p. 25) as for mappings of emotional valence, number, size, and a host of other dimensions. Our findings suggest, however, that it may be premature to exclude emotional intensity from this group. The results of Experiment 2 show a congruity effect for emotional intensity similar to effects regarded as evidence for spatialization of other dimensions (e.g., Dehaene et al., 1993; Fabbri et al., 2012; Ren et al., 2011), including P&C's for emotional valence. Although many such effects may be explained by polarity correspondence, other evidence for spatialization of number and emotional valence cannot (Casasanto, 2014; Santiago & Lakens, 2015)—suggesting that emotional intensity might follow a similar pattern. Indeed, H&L's study was motivated by the possibility that different magnitudes—numerical, temporal, spatial, emotional, and perhaps others (cf. Macnamara et al., 2019)—may be spatialized similarly, and it is unclear what theory would predict that emotional intensity should deviate from the rest. The present findings are thus in line with H&L's original proposal, laying the foundation for further exploration of similarities (as well as possible differences) in how emotional intensity and other dimensions are mapped to left–right space.

Acknowledgments

We thank Ian Carey, Evan Doherty, Madeleine Garcia, Andrew Kirvin-Quamme, and Frances Murray for help with data collection and coding. We are also grateful to Daniel Casasanto and two anonymous reviewers for their helpful feedback on a previous version of this manuscript.

Note

1. We thank Daniel Casasanto for helpful discussion of these possibilities.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bonn, C. D., & Cantlon, J. F. (2017). Spontaneous, modality-general abstraction of a ratio scale. *Cognition*, *169*, 36–45.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*, 1831–1840.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*, 83–91.

Casasanto, D. (2014). Bodily relativity. In L. Shapiro (Ed.), *Routledge handbook of embodied cognition* (pp. 108–117). New York: Routledge.

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*, 132–147.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371–396.

Fabbri, M., Cancellieri, J., & Natale, V. (2012). The A Theory of Magnitude (ATOM) model in temporal perception and reproduction tasks. *Acta Psychologica*, *139*, 111–123.

Fias, W., & Fischer, M. H. (2005). Spatial representation of numbers. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 43–54). New York: Psychology Press.

Harvey, B. M., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences*, *112*, 13525–13530.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*, 389–395.

Holmes, K. J., & Lourenco, S. F. (2011). Common spatial organization of number and emotional expression: A mental magnitude line. *Brain and Cognition*, *77*, 315–322.

Holmes, K. J., & Lourenco, S. F. (2013). When numbers get heavy: Is the mental number line exclusively numerical? *PLoS ONE*, *8*, e58381.

Lourenco, S. F. (2016). How do humans represent numerical and nonnumerical magnitudes? Evidence for an integrated system of magnitude representation across development. In A. Henik (Ed.), *Continuous issues in numerical cognition: How many or how much* (pp. 375–403). London, UK: Elsevier Academic Press.

Lourenco, S. F., Ayzenberg, V., & Lyu, J. (2016). A general magnitude system in human adults: Evidence from a subliminal priming paradigm. *Cortex*, *81*, 93–103.

Macnamara, A., Loetscher, T., & Keage, H. A. D. (2019). Judging risk magnitude: Walking to the left and base jumping to the right. *Experimental Brain Research*, *237*, 653–662.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinbugh inventory. *Neuropsychologia*, *9*, 97–113.

Pitt, B., & Casasanto, D. (2018). Spatializing emotion: No evidence for a domain-general magnitude system. *Cognitive Science*, *42*, 2150–2180.

Proctor, R. W., & Xiong, A. (2015). Polarity correspondence as a general compatibility principle. *Current Directions in Psychological Science*, *24*, 446–451.

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/. Accessed May 6, 2018.

Ren, P., Nicholls, M. E., Ma, Y. Y., & Chen, L. (2011). Size matters: Non-numerical magnitude affects the spatial coding of response. *PLoS ONE*, *6*, e23553.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, *57*, 493–502.

Santiago, J., & Lakens, D. (2015). Can conceptual congruency effects between number, time, and space be accounted for by polarity correspondence? *Acta Psychologica*, *156*, 179–191.

Sellaro, R., Treccani, B., Job, R., & Cubelli, R. (2014). Spatial coding of object typical size: Evidence for a SNARC-like effect. *Psychological Research*, *79*, 950–962.

Song, X., Chen, J., & Proctor, R. W. (2017). Role of hand dominance in mapping preferences for emotional-valence words to keypress responses. *Acta Psychologica*, *180*, 33–39.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153–181.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249.

Vallesi, A., Binns, M. A., & Shallice, T. (2008). An effect of spatial-temporal association of response codes: Understanding the cognitive representations of time. *Cognition*, *107*, 501–527.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488.

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science*, *50*, 489–525.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Appendix S1. Participant-specific slope analyses.