Atten Percept Psychophys (2012) 74:1226–1240
DOI 10.3758/s13414-012-0297-9

Size–brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning

Peter Walker · Laura Walker

Published online: 7 April 2012

Abstract Following Karwoski, Odbert, and Osgood (Journal of General Psychology 26:199–222, 1942), in the present article, cross-sensory correspondences are proposed to reflect the alignment of, and extensive bidirectional cross-activation among, dimensions of connotative meaning. The size–brightness correspondence predicted on this basis (in which smaller is aligned with brighter) was confirmed in two ways. First, when participants explored three wooden balls of different size by touch alone and indicated how bright they thought each of them was, the smaller ball was judged to be brighter than the bigger ball. Second, when these two balls served as response keys in a speeded brightness-classification task, participants were quicker and more likely to be correct when confirming that a stimulus was bright (dark) when this required them to press the smaller (bigger) key, than when it required them to press the bigger (smaller) key. This congruity effect originated from interactions embedded in the later stages of information processing concerned with stimulus classification and response selection. These results, together with the observation that the cross-sensory features associated with smallness are the same as those associated with higher pitch sounds (i.e., both attributes are more active, brighter, faster, lighter in weight, quieter, sharper, and weaker than their opposites), support the suggestion that there exists a core set of cross-sensory correspondences that emerges whichever stimulus feature is used to probe it.

P. Walker (✉) · L. Walker
Department of Psychology, Lancaster University,
Lancaster LA1 4YF, UK
e-mail: p.walker@lancaster.ac.uk

Keywords Size-brightness correspondence · Cross-sensory correspondences · Connotative meaning · Congruity · Speeded classification

Our sensory systems provide different types of converging evidence about objects and events. For example, they provide *equivalent* information (see Marks, 1978) about the same measurable feature of an object, as when vision and touch both provide information about its surface texture. Different sensory channels also provide *corresponding* information (what Marks, 1978, calls *analogous* information) about an object, as when vision and audition both indicate its connotative brightness, the former by registering its surface brightness, the latter by registering that it makes relatively *bright* (i.e., high-pitched) sounds.[1] The provision of corresponding connotative information through different sensory channels was the focus of the present study.[2]

Early evidence for cross-sensory correspondences has come from studies of sound symbolism (e.g., Kohler, 1929; Sapir, 1929), visual-hearing synaesthesia (see Marks, 1975, for a review), and the universal connotative meanings of elementary stimulus features (e.g., Karwoski, Odbert, & Osgood, 1942). In visual-hearing synaesthesia, for example, higher pitched sounds induce visual forms that are brighter,

[1] The connotations (connotative meaning) of any kind of stimulus are what it suggests, implies, or invokes, rather than what it explicitly or directly denotes.

[2] Sensory channels are rather loosely conceived to be those parts of the sensory systems responsible for encoding dissociable stimulus features, such as color, texture, and movement in vision. The term *sensory channel* is preferred over *sensory modality* because correspondences exist among feature dimensions within the same modality (e.g., brightness and angularity/pointedness in vision). For this reason, the term *cross-sensory correspondences* is preferred over *cross-modality correspondences* (e.g., Spence, 2011).

higher in space (i.e., more elevated), more likely to be moving, sharper, and smaller than those induced by lower-pitched sounds (e.g., Karwoski & Odbert, 1938; Marks, 1978; Ward, Huckstep, & Tsakanikos, 2006; Zigler, 1930). The same correspondences have since been confirmed in nonsynaesthetes who, for example, judge higher pitched sounds to be more active, brighter, faster, higher in space, lighter in weight, quieter, sharper, smaller, and weaker than lower pitched sounds (Collier & Hubbard, 2001; Eitan & Timmers, 2010; Hubbard, 1996; Marks, 1974, 1975, 1978; Mondloch & Maurer, 2004; Perrott, Musicant, & Schwethelm, 1980; Walker & Smith, 1984). And, when asked to draw the visual imagery induced by listening to short musical selections, nonsynaesthetes draw lines and forms that tend to be thinner, brighter, smaller, and more angular (sharper), the higher in pitch and/or faster in tempo the music (Karwoski et al., 1942).

Speeded classification tasks have been especially useful for investigating cross-sensory correspondences. This is because people classify elementary stimulus features more quickly and accurately when task-irrelevant stimuli in other sensory channels have corresponding (congruent), rather than noncorresponding (incongruent), features (cf. Marks, 2004; and Spence, 2011, for reviews). For example, when classifying simple visual stimuli according to their brightness, people respond more quickly when a concurrent auditory tone has congruent pitch (i.e., high for bright, low for dark) than when it has incongruent pitch (i.e., low for bright, high for dark) (e.g., Marks, 1987).

Some researchers regard cross-sensory correspondences as an aspect of synaesthesia that is present in us all (e.g., Karwoski et al., 1942; Martino & Marks, 2001; Simner, Ward, Lanz, Jansari, Noonan, Glover & Oakley, 2005; Walker & Smith, 1984; Ward et al., 2006). In a recent tutorial review, however, Spence (2011) proposed that cross-sensory correspondences reflect three aspects of normal sensory information processing. First, among dimensions concerned with the magnitude or strength of stimuli,[3] correspondences reflect the presence of a common basis for coding stimulus magnitude (e.g., the rate of neural firing). Spence referred to this type of correspondence as being *structural* in nature, and saw the correspondence between loudness and brightness as belonging to this category. Second, Spence proposed that some correspondences reflect an acquired appreciation of natural co-occurrences among features. He suggested, for example, that the correspondence between auditory pitch and size is learned through exposure to the natural tendency for smaller objects and smaller animals to generate higher pitched sounds. Spence referred to this type of correspondence as having a *statistical* basis and included the correspondences between pitch and elevation, pitch and size, and loudness and size in this category. Finally, some correspondences are proposed to arise because the same words are used to mark contrasting values on different dimensions of sensory experience (e.g., the words *high* and *low* mark contrasting levels of pitch, spatial elevation, and spatial frequency). As a person becomes familiar with the language of their culture, correspondences are established wherever such shared terminology is encountered. Spence regarded these correspondences as being *linguistic* in nature (although *lexical* might be more appropriate), and he included the correspondences between pitch and elevation, and pitch and spatial frequency, in this category.[4]

Spence (2011) acknowledged that future research will probably reveal correspondences requiring his classification to be extended (e.g., he anticipated correspondences based on the affective connotations, or valence, of stimuli). However, there is already some evidence for a correspondence between size and brightness (specifically, surface brightness)[5] that is of a kind that is not amenable to a structural, statistical, or linguistic explanation (Walker & Smith, 1985). This size–brightness correspondence was the focus of the present study.

The various correspondences involving auditory pitch, first observed in visual-hearing synaesthesia, and then confirmed in nonsynaesthetes (see above), include its associations with size and brightness. If correspondences are functionally bidirectional (cf. Martino & Marks, 2001) and reflect extensive crosstalk between dimensions of connotative meaning (cf. Karwoski et al., 1942), then there should exist a core set of correspondences that is accessible to stimuli in any sensory channel, with correspondences

[3] The distinction between magnitude dimensions and other types of dimension of sensory experience (what Stevens, 1957, called *prothetic* and *metathetic* continua, respectively) is, in essence, a distinction between quantity and quality. According to Stevens (1957, 1975), prothetic dimensions, such as loudness, are concerned with how much of a type of sensation is being experienced and are linked to variations in the level of excitation of a sensory channel. Metathetic dimensions, such as auditory pitch, are concerned with qualitative differences in sensory experience and are linked to variations in the location of maximum excitation in a sensory channel.

[4] The term *lexical* refers to those features of a word (typically its sequence of letters) that are used to find the word's entry in a dictionary, without reference to any of its meanings. Only with this restricted view of a word can it be said, for example, that *lightness* in weight and *lightness* in color share the same word (i.e., *light*). Once word meaning is brought into play, *light* in weight and *light* in color are different words. Although Spence (2011) suggested that linguistic correspondences also become semantic in nature, it remains unclear how and why this should happen.

[5] Illuminance (the amount of light provided by a source) and surface luminance (the amount of light coming from a surface) are both often referred to as *brightness*. Although most studies of cross-sensory correspondences involve surface luminance, a few involve illuminance (see Lewkowicz & Turkewitz, 1980; Marks, 1978; Stevens, 1975). It is the former, surface luminance, here referred to as *brightness* or *surface brightness,* which was of concern in the present study.

demonstrating some kind of transitivity in relation to the direction in which they align themselves. On this basis, a correspondence between size and brightness is to be expected. Thus, if smaller is aligned with higher pitch, and higher pitch is aligned with brighter, then smaller should be aligned with brighter (see Marks', 1978, account of Von Hornbostel, 1931, demonstration of the transitivity of the correspondences involving surface brightness, odor, and auditory pitch).

In what ways would a size–brightness correspondence of this nature fall outside of Spence's (2011) framework? If surface brightness is not distinct from illuminant brightness and, like the latter, behaves as a magnitude dimension, then the nature of the predicted correspondence would contradict a structural account, because on a structural account, smaller (*less* size) should be aligned with darker (*less* light), rather than with brighter (*more* light). Of course, if surface brightness is distinct from illuminant brightness and does not behave as a magnitude dimension, as the evidence suggests (see, e.g., Marks, 1974, 1982, 1987; Smith & Sera, 1992; Wicker, 1968; see the Appendix for an account of the evidence), then this in itself precludes it from entering into a correspondence with a structural basis.

With regard to the predicted size–brightness correspondence having either a statistical or a linguistic basis, there is little evidence to support either possibility. Although it is acknowledged that without a comprehensive analysis of either the natural world of objects, or of the language of cross-sensory correspondences, it is difficult to make claims with any confidence, it has been claimed that there are no natural co-occurrences from which the correspondence could be derived (Mondloch & Maurer, 2004; Smith & Sera, 1992), and it is difficult to find any words in common use (at least in English) that mark contrasting values on both dimensions. Of course, even if future research finds some relevant co-occurrences, additional evidence would still be required to demonstrate that the correspondence is acquired through exposure to them. Therefore, for the moment at least, it seems that any demonstration of the predicted size–brightness correspondence would encourage the extension of Spence's (2011) classification of correspondences, perhaps to include a type of correspondence based on the connotative meanings of elementary stimulus features, as Karwoski et al. (1942) had proposed. A main aim of the present study was to provide such a demonstration.

As was indicated already, there is some evidence, albeit very modest in nature, for a size–brightness correspondence in which smaller is aligned with brighter. In a study in which participants indicated which of two objects contrasting on a single feature dimension (e.g., a big and small version of a toy mouse) matched an object with an extreme feature on another dimension (e.g., a medium-sized toy mouse that was either black or white), Smith and Sera (1992) observed the predicted correspondence in 2-year-old children, but failed to observe it consistently in older children and adults. However, when Walker and Smith (1985) asked adults to explore wooden balls varying in size by touch alone, and to comment on the attributes each possessed, they did observe a consistent trend whereby smaller balls were judged to be brighter than larger balls.

In the present context, it is relevant to note that Walker and Smith (1985) went on to demonstrate that, as a task-irrelevant stimulus feature, haptic size can give rise to a congruity effect in a speeded classification task. They did so in a way that revealed something about the basis of the correspondence responsible for the congruity effect. A spatial elevation classification task was designed around the correspondence between elevation and size (in which smaller is aligned with higher), and contrasting values for elevation were presented verbally. Participants were shown a single word selected at random from two pairs of polar adjectives linked to opposite ends of the spatial elevation dimension (i.e., TOP–BOTTOM and UP–DOWN). The two adjectives associated with the same end of this dimension (i.e., TOP–UP and BOTTOM–DOWN) were assigned to the same response key (i.e., left or right). The two response keys, which were never seen by participants, differed in size, although not in their haptically perceived height. Which size of key was assigned to which hand was surreptitiously switched across successive blocks of trials so that, as an incidental feature of the task situation, participants sometimes responded on the small key to the words TOP and UP, and on the big key to the words BOTTOM and DOWN (congruent condition), and sometimes responded on the big key to the words TOP and UP, and on the small key to the words BOTTOM and DOWN (incongruent condition). Participants responded more quickly and accurately in the congruent condition, confirming their sensitivity to the correspondence between (haptic) size and spatial elevation.

Cross-sensory correspondences have the potential to give rise to congruity effects at various stages of information processing, which Spence (2011) broadly categorised as early (sensory) and late (decisional) stages.[6] With this in mind, it is important to acknowledge two features of Walker and Smith's (1985) study implicating interactions at

[6] Spence (2011) proposed that the level at which congruity effects arise is related to the type of correspondence responsible for the effects. He suggests that, in general, congruity effects based on correspondences having a structural or statistical basis reflect interactions at relatively low (early) levels of information processing, allowing them to have an impact on perception. On the other hand, congruity effects based on correspondences having a linguistic basis are thought to reflect interactions at higher levels of processing, such as those involved in stimulus classification and response selection (i.e., decisional processes).

relatively late stages. First, and most obvious, is Walker and Smith's (1985) decision to use words to specify the contrasting levels of elevation being classified. The second, less obvious feature is the manner in which haptic size was introduced to the task situation. Participants grasped the response keys continuously throughout a block of trials, and it was only when a test word had been classified that one of the keys, and one of the two contrasting sizes of key, became salient. Up to that point it was not known which key would need to be pressed, and so both keys, and both sizes of key, remained equally salient. The congruity effect must, therefore, have originated in relatively late stages of information processing, such as those concerned with stimulus classification and response selection (see the General Discussion section below).

This second feature of Walker and Smith's (1985) study is independent of the verbal or nonverbal nature of the stimuli being classified. It should be possible, therefore, to use nonverbal stimuli to demonstrate a size–brightness congruity effect at the same time as exploiting haptic size to isolate relatively late stages of information processing as the basis for the effect. Providing such a demonstration was an additional aim of the present study.

To summarize, in the main present experiment, a speeded brightness classification task was employed in which nonverbal visual stimuli were to be classified on the basis of their surface brightness, and in which participants registered their decisions by pressing the left or right of two keys which, incidentally, contrasted in size. The instructions specified which key should be pressed for each level of brightness by referring to the *left* and *right* key. The difference in size of the two keys was task irrelevant and was never mentioned by the experimenter. Nevertheless, it was expected that a size–brightness correspondence, in which smaller is aligned with brighter, would give rise to a congruity effect, with participants responding more quickly and accurately when key size and brightness were in correspondence.

Before embarking on the experiment, however, the cross-sensory features possessed by the different sized objects that were to serve as response keys needed to be confirmed, especially that the smaller object is judged to be brighter when explored by touch alone. In addition to providing an alternative way of confirming the predicted size–brightness correspondence, having the objects rated for their cross-sensory features offered an opportunity to see whether the features associated with size were those predicted on the basis of the existence of a core set of correspondences. This core set of correspondences was provisionally identified as those emerging when auditory pitch has been examined in previous studies (e.g., Eitan & Timmers, 2010; Walker & Smith, 1984).

Preliminary observations: cross-sensory connotations of haptic size

Method

Participants

Fifty-two Lancaster University students (nine males and 43 females, mean age = 18.9 years, age range = 18 to 28 years), none of whom were tested for synaesthesia, volunteered to participate in exchange for course credit or payment. All of the participants in the present study were recruited using the Sona Systems experiment management system, with no requirements specified regarding gender or handedness, and with no mention of synaesthesia. All but two of the participants in this preliminary study were right-handed by self-report, and 28 of them had already completed the main experiment reported below (i.e., the speeded brightness classification task) before providing the ratings reported in this section.[7]

Materials

Participants rated three smoothed wooden balls that they explored by touch alone on each of nine feature dimensions. The smallest and largest balls were 2.5 and 7.5 cm in diameter, respectively, and matched those to be used later as the small and large response key in the speeded classification task. The ball of intermediate size was 5 cm in diameter. To ensure that all three balls were perceived (haptically) to be equally high in space, the medium and small ball were raised from the table, on small wooden blocks, by 2 and 3.75 cm, respectively. The balls were always covered with a thick black cloth and were never seen by participants.

Two of the cross-sensory feature scales on which each object was to be rated were anchored at their endpoints with the antonyms *bright–dark* and *high–low (in space)*. With regard to height, it was important to ensure that the perceived height of the balls was not confounded with their size, thereby isolating size as the only objective feature distinguishing the two response keys in the speeded classification task. Six additional cross-sensory feature scales were introduced to assess the extent to which the set of correspondences observed with auditory pitch would reemerge when variations in the size of an object were being considered. These scales were anchored by the antonyms *active–passive, fast–slow, high–low (pitch)*, *quiet–*

[7] It was arranged for these participants to provide ratings after, rather than before, completing the brightness classification task to avoid contaminating their performance in this task. The lack of contamination was confirmed by entering involvement, or not, in the rating task as a between-participants factor in the analysis of performance in the classification task. This factor had no impact on classification performance. Equally, the responses in the rating task were uninfluenced by whether participants had previously completed the classification task.

loud, sharp–blunt, and *weak–strong*. Finally, because the affective value of a stimulus has the potential to provide a generic basis for cross-sensory correspondences, a scale assessing the valence of each value of size was included. This scale was anchored by the antonyms *bad–good*.

The scales on which the balls were to be rated were represented visually with a horizontal line along which there were six, equally spaced, vertical marks. Positioned immediately below successively less extreme marks were the verbal labels *very, quite*, and *slightly*. Participants were asked to circle the mark which best represented a ball's position on the scale. An integer score of 1–6 was recorded, with 1 and 6 being assigned to the two extreme *very* ratings. A neutral midpoint position (i.e., a *neither* rating) was not available to participants, although an average score of 3.5 across participants would indicate such neutrality for the group.

Procedure

Before providing any ratings, participants familiarized themselves with the three objects by exploring them haptically. The objects were arranged on a table in a horizontal row with their left–right positioning determined at random, afresh for each participant. On each trial, participants were first instructed which ball (labeled *A*, *B*, and *C*) to explore and were then presented with the rating scale. Each of the three balls was rated once on each of the nine scales, resulting in a total of 27 trials. The presentation order of the objects and scales was determined at random, afresh for each participant.

Results

Table 1 gives the mean rating for each of the three balls on each of the scales, along with statistical confirmation (using sign tests) regarding which ball's ratings were significantly above or below the neutral midpoint of each scale (i.e., a value of 3.5). The ratings for each scale were also submitted to Page's (1963) L trend test to assess the significance of any systematic changes with increasing ball size. Table 1 indicates for which scales there was a significant trend in the ratings across the three values for size. An alpha level of .05 was adopted in all statistical analyses reported in the present study.

For none of the balls was the mean rating for spatial height significantly different from the midpoint on the scale, neither was there a significant trend across levels of ball size. These null outcomes confirm that raising the balls above the level of the table according to their size had been successful in ensuring they were perceived (haptically) to be equally high. Ball size was associated with a significant trend on all of the other scales, except *bad–good*. With regard to this valence scale, inspection of the individual trend statistic across participants failed to reveal a bimodal distribution. Hence, the lack of a significant trend for the group as a whole did not arise because a subset of participants were seeing small as good, whereas the others were seeing it as bad.

Discussion

The ratings consolidate and extend Walker and Smith's (1985) observation that smaller haptic size is aligned with brighter, thereby confirming the predicted size–brightness correspondence. More specifically, whereas the mean brightness rating for the smallest ball was midway between *slightly bright* and *quite bright*, the mean rating for the largest ball was approximately midway between *slightly dark* and *quite dark*. Importantly, the ratings also confirm the prediction that the cross-sensory features associated with haptic size will be the

Table 1 Mean ratings (*SEM*s in parentheses) for each size of ball on nine 6-point scales. Mean ratings below 3.5 indicate that the antonym listed first for each scale was judged more appropriate

Rating Scales	Ball Size						
	Small		Medium		Large		
	M	*SEM*	*M*	*SEM*	*M*	*SEM*	Trend
High–low (space)	3.90	0.27	3.45	0.15	3.35	0.27	$L = 608$
Active–passive	2.34	0.20***	3.16	0.18*	4.13	0.22**	$L = 679$ ***
Bright-dark	2.54	0.22***	3.45	0.15	4.39	0.22***	$L = 678$ ***
Fast-slow	2.10	0.22***	3.28	0.15*	4.48	0.22***	$L = 684$ ***
High–low (pitch)	2.01	0.20***	3.70	0.14	5.33	0.15***	$L = 705$ ***
Quiet–loud	2.59	0.23***	3.73	0.15	4.88	0.21***	$L = 682$ ***
Sharp–blunt	3.63	0.26	4.82	0.15***	5.45	0.12***	$L = 680$ ***
Weak–strong	2.53	0.22***	4.36	0.13***	5.40	0.14***	$L = 698$ ***
Bad–good	4.00	0.20*	4.00	0.11***	3.67	0.22	$L = 614$

*** $p < .001$, ** $p < .01$, and * $p < .05$ for mean ratings that are significantly different from the neutral value of 3.5 by sign test, and for significant trends across levels of size using *Page's L* trend test

same as those associated with auditory pitch (i.e., both smallness and high pitch are more active, brighter, faster, lighter in weight, quieter, sharper, and weaker than their opposites).

The presence of this size–brightness correspondence is consistent with the proposed functional transitivity among correspondences (on which basis the correspondence was predicted) and with their bidirectionality. Thus, whereas previous studies have revealed that high-pitched sounds are small (see the Introduction), the present results show that small objects are high pitched. Transitivity and bidirectionality together imply that the same core set of correspondences should emerge regardless of the stimulus feature used to probe it, and this is entirely consistent with the proposal that the correspondences reflect interactions among dimensions of connotative meaning (cf. Karwoski et al., 1942).[8] For example, because the same amodal connotative features are accessed by auditory pitch and haptic size, the same set of correspondences should be observed, whichever of these features is used to probe them (see Fig. 2).

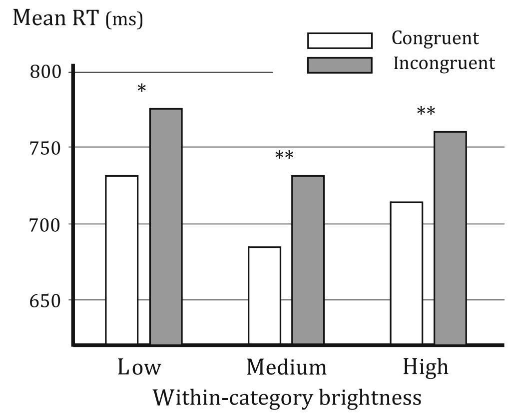

Fig. 1 Experiment: Mean *RT*s for congruent and incongruent combinations of categorical brightness and key size at each level of within-category brightness. * $p < .01$; ** $p < .001$

Experiment: size–brightness congruity in speeded classification

In the main experiment, participants performed a speeded brightness-classification task in which they were required to decide whether a visual stimulus was brighter or darker than the midgray background against which it appeared (cf. Marks, 1987, for confirmation that manipulating brightness in this way isolates surface brightness from illuminant brightness). They indicated their decision by pressing one of two keys that differed in size. It was assumed that a cross-sensory correspondence between size and brightness exists and that, like other correspondences, it can have an impact on performance as an incidental feature of a task situation. Accordingly, it was expected that participants would respond more quickly and with fewer errors when the size of the response key needing to be pressed was congruent with the brightness classification of the visual stimulus (i.e., *bright*–small, *dark*–big), rather than when it was incongruent with it (i.e., *bright*–big, *dark*–small).

Method

Participants

Seventy-six Lancaster University students (62 females and 14 males) between the ages of 18 and 34 (mean age = 19.75 years), none of whom were tested for synaesthesia, volunteered to participate for course credit or payment. All but four of the participants were right-handed by self-report.

Materials

The visual stimuli were six solid circles that were 4.5 cm in diameter, varying in brightness from white through to black. The circles were presented individually at the center of a 20-in. computer screen (Apple A1038, 1,680 × 1,050 cinema back-lit LCD display; response time [RT] = 14 ms, controlled by a Dual 2 GHz, PowerMac G5) and were brighter (340, 230, & 150 cd/m^2) or darker (42, 17, and 2 cd/m^2) than the midgray (90 cd/m^2) background against which they appeared. The response keys comprised the smallest (2.5-cm diameter) and largest (7.5-cm diameter) of the wooden balls whose attributes were assessed in the preliminary observations. These were mounted on microswitches, whose physical resistances were adjusted until the two authors judged that equal force was needed to close each of them. This required a higher level of resistance to be set for the big key (i.e., 1,000 gm) than for the small key (250 gm). The small key was also raised 3.75 cm from the table to ensure that the two balls were perceived (haptically) to be equally high spatially (see the Preliminary Observations section). The response keys were covered with a thick black cloth throughout the experiment, ensuring that participants were never able to see them.

Design and procedure

Participants were instructed to grasp the top half of each key with the thumb and first two fingers of each hand, and to continue grasping them until told to take a break between blocks of trials. Immediately a circle appeared, participants classified it as *bright* or *dark* (relative to the midgray

[8] In the context of visual-hearing synaesthesia, Karwoski et al. (1942) proposed that "The synesthetic or analogical process appears to be the parallel alignment of two gradients in such a way that the appropriate extremes are related, followed in some cases by translation in terms of equivalent parts of the two gradients thus paralleled" (p. 217).

background) by pressing one of the two keys. Half of the participants (30 females and eight males) were asked to press the left-hand key when the circle was brighter than the background and the right-hand key when the circle was darker than the background. The other participants (32 females and six males) were assigned to the opposite brightness-hand mapping (i.e., right for *bright*, left for *dark*). The difference in the sizes of the two keys was incidental to the speeded-classification task and was never mentioned by the experimenter.

Participants completed four blocks of trials, each separated by a 1-min break. In a block of 48 trials, each of the six circles appeared eight times, in a randomly determined order that was generated afresh for each participant. Each circle remained visible until participants made their brightness decision and was followed by a blank interval of 3 s before the next circle was presented. Participants did not receive feedback about the speed or correctness of any of their responses.

At the end of each trial block, the experimenter surreptitiously switched the left–right positions of the two response keys so that participants performed the proceeding block using the opposite mapping of brightness on to key size. Thus, across the four experimental blocks, participants alternately pressed the small key for *bright* and the big key for *dark* (congruent condition), or the small key for *dark* and the big key for *bright* (incongruent condition). We counterbalanced across participants which of these two mappings (small vs. big for *bright*) was assigned to the first block of trials.

Results

Probability correct

Although the overall level of correct responding was consistently high, it was significantly higher for congruent trials, p(correct) = .99, than for incongruent trials, p(correct) = .98, Wilcoxon Signed Ranks $p < .001$.

Reaction time

RTs from incorrect trials were excluded from the analysis, and RTs greater than 2.5 *SD*s above a participant's overall mean correct RT were replaced with the cut-off value. The resulting overall mean correct RT was 733 ms. Mean correct RTs for the various conditions are given in Table 2.

Two aspects of the variation in surface brightness were treated as separate factors in an ANOVA. The first, between-category brightness, refers to the brightness feature on which response selection was based and relates to whether a circle was brighter or darker than the midgray background. The second, within-category brightness, refers to the variation in surface brightness within each level of categorical brightness. Thus, for each level of categorical brightness, within-category brightness was low, medium, or high. Where a circle was darker than the background, these levels of within-category brightness were 2 (black), 17, and 42 cd/m^2, respectively. Where a circle was brighter than the background, the low, medium, and high values for within-category brightness were 150, 230, and 340 (white) cd/m^2, respectively.

Mean correct RTs (i.e., means of eight individual RTs) were submitted to a 2 × 2 × 3 × 2 mixed ANOVA, with key size (small vs. big), between-category brightness (darker vs. brighter), and within-category brightness (low, medium, or high) as within-participant factors. Hand-brightness assignment (left- vs. right-hand key for *bright*) was a between-participants factor.[9]

Although there was no main effect of key size, $F(1, 74) = 2.67$, $p = .11$, $\eta_p^2 = .04$, (mean *RT*s for the small and big key being 724 ms and 743 ms, respectively), there was a significant main effect of between-category brightness, $F(1, 74) = 17.32$, $p < .001$, $\eta_p^2 = .19$, with participants classifying darker circles more quickly (mean RT = 714 ms) than brighter circles (mean RT = 752 ms). There was a significant between-category brightness × key size interaction, $F(1, 74) = 21.56$, $p < .001$, $\eta_p^2 = .23$, the nature of which confirmed the predicted congruity effect. The mean RTs for congruent and incongruent trials were 710 and 756 ms, respectively.

There was a main effect of within-category brightness, $F(2, 148) = 27.72$, $p < .001$, $\eta_p^2 = .27$, largely attributable to a significant quadratic component, $F(1, 74) = 47.64$, $p < .001$, $\eta_p^2 = .39$. Within-category brightness interacted significantly with between-category brightness, $F(2, 148) = 140.36$, $p < .001$, $\eta_p^2 = .66$, and the nature of this interaction confirmed that participants responded more quickly the higher the brightness contrast (in either direction) between the circle and the midgray background. Figure 1 shows that, despite the very pronounced effects of within-category brightness, and the marked changes in the effect of between-category brightness across different levels of within-category brightness (see Table 2), the difference in response speed between congruent and incongruent trials remained within 40–50 ms across successive levels of local brightness (i.e., 45.5, 49.5, and 44.5 ms, across low, medium, and high levels of within-category brightness, respectively). The size–brightness congruity effect appears here to be very robust.

There was not a significant interaction between within-category brightness and key size, $F < 1$, indicating that

[9] A preliminary analysis of the results always explored the potential influence of the gender of participants. Here, as in other of our studies on cross-sensory correspondences, gender did not moderate the effects of any of the factors of interest here. For this reason, it was not included as a factor in the ANOVA reported here.

Table 2 Experiment: Mean correct RTs (*SEMs* in parentheses) and *p* (correct) according to the between-category and within-category brightness levels, and key size, with associated differences in RT attributable to key size

Brightness:						
Between Category	Dark			Bright		
Within Category	Low	Medium	High	Low	Medium	High
Key Size						
Small	676	705	803	**814**	**689**	**655**
	(*18*)	(*21*)	(*26*)	**(*28*)**	**(*19*)**	**(*17*)**
	.99	.98	.98	**.98**	**.99**	**1.00**
Big	**649**	**678**	**774**	878	761	715
	(*18*)	**(*19*)**	**(*24*)**	(*33*)	(*28*)	(*22*)
	1.00	**1.00**	**.98**	.97	.99	.99
Difference (RT_{big}–RT_{small})						
	−27	-27	-29	64	72	60

Bold entries relate to congruent conditions; normal font entries relate to incongruent conditions

variations in this aspect of brightness (i.e., variations within a level of categorical brightness) did not combine with key size to generate a congruity effect. The lack of a significant three-way interaction between within-category brightness, between-category brightness, and key size, $F < 1$, indicates that the level of brightness contrast did not interact with key size.

There was not a significant main effect of hand-brightness assignment, $F < 1$, and no significant interaction between this and any other factor, or combination of factors (all *p*s > .05). No other interactions were significant.

Discussion

When a small and a large ball were used as response keys in a speeded brightness-classification task, a congruity effect was observed confirming the predicted correspondence between size and brightness. Participants responded more quickly and with fewer errors when the relative size of the key needing to be pressed was congruent with the relative brightness of the visual stimulus being classified. That is, they responded relatively more quickly and accurately when pressing the small (big) key to confirm that a circle was bright (dark), than when pressing the small (big) key to confirm that a circle was dark (bright). This happened even though key size was task irrelevant. Before accepting the results as evidence for the predicted size–brightness correspondence, however, two factors supporting alternative explanations need to be assessed.

Brightness contrast

The potential for brightness contrast to interact with key size to create a congruity effect must be considered because brightness contrast is known to enter into a generic correspondence with other features according to their magnitude (cf. Walsh, 2003). For example, people find it easier to encode numerically bigger (smaller) digits when they are presented at higher (lower) levels of visual contrast relative to the background on which they appear (R. Cohen Kadosh & Henik, 2006; Cohen Kadosh, Cohen Kadosh & Henik, 2008).[10] However, although the level of brightness contrast in the present experiment had a very significant effect on performance, with participants generally responding more quickly with higher levels of contrast, it did not interact with key size to produce a congruity effect. Thus, participants did not respond quicker when *more* contrast was paired with the big key (*more* size) than when it was paired with the small key *(less* size). Neither did they respond quicker when *less* contrast was paired with the small key, than when it was paired with the big key. Interestingly, in a situation in which there is no background to take into account, digit-color synaethetes associate brighter colors with numerically smaller digits (Cohen Kadosh, Henik, & Walsh, 2007), reinforcing the alignment on which the present congruity manipulation was based.

Valence as a basis for the size–brightness correspondence

Spence (2011) acknowledged that his framework might need to be extended to embrace a generic correspondence based on the valence of stimuli (i.e., evaluation or affective potential). Karwoski et al. (1942) also considered valence to be a possible basis for cross-sensory correspondences, and Proctor and Cho (2006) discussed the pervasiveness of valence-based congruity effects in speeded binary classification tasks. With size and brightness both being regarded as valenced dimensions, so that *big* (Meier, Robinson, &

[10] Comparing these two studies makes it clear that it is the contrast in surface brightness that contributes to this correspondence effect, rather than the surface brightness of the digits, per se.

Caven, 2008; Silvera, Josephs, & Giesler, 2002) and *bright* (Meier, Robinson & Clore, 2004; Meier, Robinson, Crawford, & Ahlvers, 2007) are both deemed, by some, to be positively valenced, a valence-based explanation of the size–brightness congruity effect needs to be assessed.

According to one valence account, participants should find it easier to respond when the brightness of a circle and the size of the key being pressed have the same valence. However, if *big* and *bright* are both positive, this valence account predicts the reverse of the congruity effect observed in the present study. Thus, participants should have responded more quickly and accurately when the big (+) key was assigned to bright (+) and the small (−) key to dark (−). This was not the case.

According to a second valence account, it is the valence associated with hand dominance that plays the critical role in providing an alternative explanation for the congruity effect observed presently. According to Casasanto (2009), the dominant and nondominant hands are positively and negatively valenced, respectively. If the valence correspondence between brightness and hand contributed to the congruity effect, however, participants should have shown better overall performance when *bright* (+) was assigned to the dominant (+) right hand, than to the nondominant (−) left hand. In addition, the impact of size–brightness correspondence (i.e., the congruity effect) should have been more pronounced in the former case. Neither of these effects was observed.

These arguments are based on the assumption that big is always positively valenced. However, the valence associated with size is highly context sensitive. Although a big apple, relative to a small apple, is likely to be positive, a big tumor is undoubtedly more negative than a small tumor. It is clear, therefore, that context is important and that it is the valence of the keys used in the main experiment that is crucial. With this in mind, it will be noted that the preliminary observations reported above failed to reveal an association between key size and valence (see Table 1). In addition, and as a final step toward eliminating valence correspondence as an explanation of the size–brightness congruity effect, one of the authors (L.W.) completed a pilot study in which she repeated the speeded classification experiment, with 24 participants classifying each of six words (from the three antonym pairs *good–bad, nice–nasty, amusing–boring*) according to their valence. There was no congruity effect involving the valence of a test word and key size, $F < 1$.

Surface brightness or illuminant brightness?

Simple images on a computer screen, such as the solid circles presented in the brightness classification task, can have a rather ambiguous status. On the one hand, they are sources of illumination, with high levels of illuminance when white, and low levels of illuminance when black. As sources of illumination, the circles would be perceived as holes in a screen (in the present study, a midgray screen), with the source of light being located in the space behind the screen. On the other hand, they are depictions of objects varying in surface brightness, and this is how they have been conceptualized in the discussion thus far. However, is there anything to confirm this interpretation, apart from the fact that the experimenters and participants seem to have perceived them as objects, and the fact that much other research in visual perception would need to be revisited if, instead, they were perceived as holes?

Brightness as illuminance behaves as a magnitude dimension, thereby entering into correspondence with a whole range of other magnitude-based dimensions, including loudness and hand-grip force (see Stevens, 1975, and Marks, 1978, for reviews). And, with demonstrations that a brighter back-illuminated aperture is perceived to be bigger than a darker aperture (Robinson, 1954), these correspondences appear to include size. It is clear, therefore, that brightness as illuminance should, and does, align itself with big, rather than with small. But this is not what was observed in the present study. Here, brightness was aligned with small, which agrees with other findings in which surface brightness has been isolated from illuminant brightness. For example, when participants in an additional rating study (Walker & Walker, 2012) indicated what cross-sensory features were possessed by circles similar to those presented in the classification task reported presently, they reported brighter circles to be smaller, weaker, quieter, and lighter in weight than darker circles, which is the opposite of what would be expected for illuminant brightness. In addition, when Marks (1987) explored congruity effects in speeded classification that involved brightness, he presented contrasting levels of brightness in each of two ways. First, he presented small rectangles varying in brightness in the center of a computer screen that was otherwise dark (i.e., black), and that remained so throughout the experiment. Second, he presented similar rectangles on a midgray background and arranged for them to be brighter or darker than the background, which is the same type of manipulation that was employed in the present study. Marks (1987) regarded the first manipulation as varying illuminant brightness (with the stimuli appearing as "luminous spots," p. 385), but regarded the second manipulation as varying surface brightness (with the stimuli appearing as "dark or light surfaces", p. 385). His interpretation was confirmed by the different patterns of congruity effect he observed for the two types of brightness manipulation. Only illuminant brightness interacted with loudness to yield a magnitude-based congruity effect (i.e., people responded more quickly and accurately when brighter visual stimuli appeared with

louder sounds) (see the Appendix, below, for further information regarding Marks, 1987 study). In summary, therefore, the weight of evidence indicates that the circles presented in the brightness classification task were perceived as objects, rather than as holes, and that the variation in brightness was encoded as variation in their surface properties.

General discussion

When participants (adult nonsynaesthetes) explored, by touch alone, three wooden balls varying in size, they judged smaller to be bright, and bigger to be dark. This correspondence between size and brightness was anticipated because both features share the same relationship with auditory pitch (i.e., higher pitched sounds are both smaller and brighter than lower pitched sounds), and because it was predicted that a kind of transitivity would hold among correspondences in terms of the directions in which dimensions align themselves (see Stevens, 1975, p. 105). Observing the predicted size–brightness correspondence is consistent with this prediction.

Participants' judgements confirmed the functional bidirectionality of the correspondence between pitch and size (i.e., high pitch is small, and small is high pitched) and, by implication, the bidirectionality of all correspondences (cf. Martino & Marks, 2001). As was pointed out already, transitivity and bidirectionality together imply that the same set of correspondences should emerge regardless of the sensory channel used to probe it. For example, with auditory pitch and haptic size accessing the same amodal connotative features, the same set of correspondences should be observed regardless of which of these features is used to probe the correspondences. Results from the preliminary study confirmed this, showing that smaller haptic objects are like higher pitched sounds (see the introduction) in being more active, brighter, faster, lighter in weight, quieter, sharper, and weaker than their opposites. This result supports the claim that there exists a core set of correspondences that is accessed, whichever sensory channel is used to probe it.

When the smallest and largest balls were used as response keys in a speeded brightness-classification task, the same size–brightness correspondence gave rise to a congruity effect. This effect occurred even though key size was task irrelevant, and one possible interpretation of this is that the size–brightness correspondence, like other correspondences, can have an impact on behavior automatically.

What kind of correspondence is that between size and brightness?

If the size–brightness correspondence does not qualify as a correspondence having a structural, statistical, or linguistic (lexical) basis (see the introduction), and does not reflect a generic correspondence based on valence, what kind of correspondence is it? That is, what new type of correspondence needs to be added to Spence's (2011) classification?

Following the lead provided by Karwoski, Osgood, and their colleagues (e.g., Karwoski et al., 1942), it is proposed that the correspondence between size and brightness reflects the manner in which the two dimensions interact at the level of connotative meaning (see Fig. 2), without linguistic mediation. Thus, haptic smallness connotes an amodal sense of smallness that then induces connotations of brightness. Similarly, visual surface brightness connotes an amodal sense of brightness that then induces connotations of smallness. It is the mutual induction of connotative meanings across dimensions (here, between connotations of size and brightness) that is the most direct basis for the size–brightness correspondence and, presumably, for some other correspondences.

Because connotative meanings are amodal in nature, they are well placed to explain the cross-sensory nature of correspondences: The same connotative meanings are induced by stimulation of any sensory channel, including the channel of most obvious relevance (e.g., connotations of brightness can be induced by visual stimulation as well as by sound). This also applies to haptic size which, as well as supporting the direct perception of size, induces connotations of size which, like other connotative meanings, are context sensitive: The same value for haptic size will induce connotations of either bigness or smallness, depending on the context in which it is encountered.

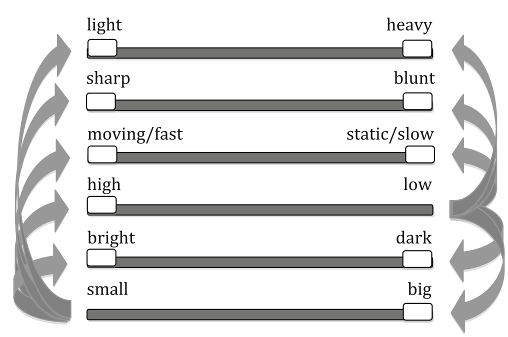

Fig. 2 Following Karwoski et al. (1942), cross-sensory associations are seen as reflecting extensive crosstalk between corresponding places on dimensions of connotative meaning. Coupled with the bidirectionality of such crosstalk, correspondences are predicted to display transitivity in terms of the directions in which the dimensions are aligned with each other. On this basis, a correspondence between size and brightness is predicted (e.g., *big* should feel *dark*). Note: Only a sample of the crosstalk between dimensions is illustrated, the point being to illustrate that the same correspondences will be in evidence whether haptic size (here, smallness) or auditory pitch (here, low pitch) is probed, implying the existence of a core set of correspondences having functional significance whichever sensory channel is used to probe the correspondences

The connotations of brightness that this, in turn, induces will vary accordingly.

The convergence of size and brightness at the level of connotative meaning appears to occur in the absence of common lexical terms marking contrasting values of size and brightness. If so, then this absence of lexical overlap distinguishes the size–brightness correspondence from the linguistic correspondences identified by Spence (2011). Perhaps, however, the latter type of correspondence also is essentially semantic in nature, with the presence of shared linguistic terms arising secondarily from this. The cross-cultural universality of the connotative meanings of elementary stimulus features (e.g., Osgood, 1960), and the correspondences they give rise to, would explain why similar linguistic terms are used to reflect the same correspondences across different languages. For example, cultures that do not use their equivalent of *high* and *low* to mark contrasting levels of auditory pitch instead use their terms for *light/heavy*, *sharp/blunt*, *small/big, thin/thick,* or *weak/strong* (cf. Eitan & Timmers, 2010). These are, of course, just the terms that the proposed set of core correspondences would predict, and their use suggests that the set of core correspondences provisionally and partially identified in the present study have cross-cultural relevance.

Semantic differential theory

The work of Karwoski et al. (1942) became situated in studies employing the semantic differential technique. A consistent finding from these studies is that connotative feature associations revolve around three independent factors identified as evaluation, potency, and activity. On this basis, therefore, at least three distinct sets of core correspondences are to be expected, encouraging a little reflection on the suggestion that only a single core set of correspondences exists.

The core set of correspondences identified in the present study do not have, as a common theme, either valence or potency. Valence was specifically ruled out as a factor in the correspondence between size and brightness. Similarly, common coding based on potency (i.e., a strength-based structural account) also was ruled out. The tripartite scheme at the heart of the semantic differential would require, therefore, that the set of correspondences revealed in the present study revolve around the concept of activity. This seems unlikely, however, not least because activity is another magnitude dimension, and yet some of the key feature dimensions within the core set of correspondences revealed here are not magnitude dimensions. Most notable among these is auditory pitch (e.g., Smith & Sera, 1992). Furthermore, Eitan and Timmers (2010) submitted their cross-sensory ratings of auditory pitch to principal component analysis, in the manner of the semantic differential technique, and although they confirmed evaluation, potency, and activity as three important underlying factors, they observed them not to be the strongest predictors of the variability among cross-sensory feature associations. This status went to a factor, which included brightness, which Eitan and Timmers found difficult to conceptualize. Just how many core sets of correspondences there are, and how they mesh with the tripartite scheme associated with studies of the semantic differential, remains for further research to clarify.

How might the congruity effect be explained?

It can be assumed that two conditions need to be in place for a particular correspondence to give rise to a congruity effect in speeded classification. First, the classification decision (and response selection) needs to refer to the same type of representation on which the correspondence is based. Second, the criterial and incidental stimulus features need to converge on this type of representation. If the size–brightness correspondence is based on connotative meanings, then the congruity effect observed in the brightness classification task should reflect interactions at relatively late (nonsensory) levels of information processing. Two features of the classification task confirm this. First, because participants grasped the two response keys continuously throughout each block of trials, one of the two keys, and its relative size, became salient only after the visual stimulus had been classified as being bright or dark. Prior to the classification of the stimulus, both keys, and both values for key size, were equally salient, and congruity was not yet an issue. Second, when a test circle was brighter or darker than the midgray background, its surface brightness could still take on any one of three values. This noticeable variation in surface brightness, which had no implications for stimulus classification (and, therefore, for response selection), did not interact with key size to yield a congruity effect. That is, within the conditions associated with a particular task-defined category of brightness (i.e., the *brighter* and *darker* conditions), participants did not respond more quickly when higher (lower) levels of surface brightness were paired with the smaller (bigger) response key. The absence of a congruity effect arising from these within-category variations in brightness is entirely consistent with the claim that the main congruity effect originated at levels of processing subsequent to the brightness classification of the stimulus.

Figure 3 illustrates the functional components determining the nature of the correspondence between size and brightness, and the basis of the congruity effect arising from this. Sensory-perceptual levels of representation are distinguished from connotative meanings, and different dimensions of connotative meaning are shown as being aligned in a way that is consistent with the core set of cross-sensory

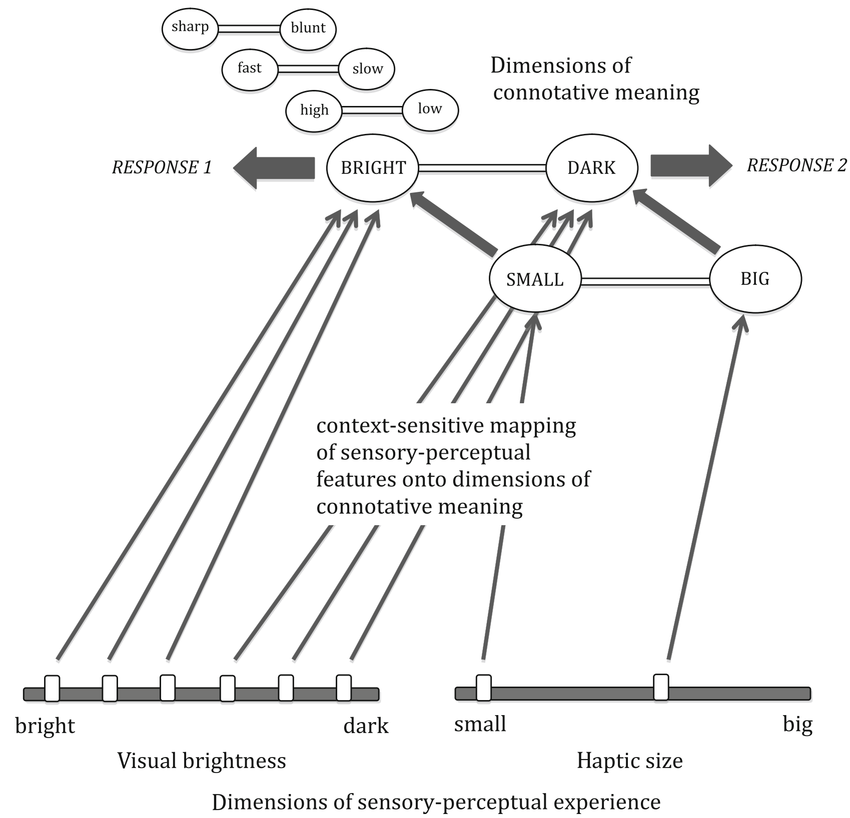

Fig. 3 The functional elements of information processing allowing the size–brightness correspondence to have an impact on performance in the brightness classification task used in the main experiment (see text for explanation)

correspondences provisionally identified here. The six different levels for visual brightness are shown as mapping onto two contrasting levels of connotative brightness, as specified in the task instructions (i.e., in a context-sensitive manner). Similarly, the two levels for haptic size are shown as mapping onto contrasting levels of connotative size, even though this mapping was not mentioned in the instructions. Although five dimensions of connotative meaning are depicted, these are not intended to be exhaustive, and the two most relevant dimensions for present purposes are emphasised. In addition, although each and every connotative dimension is assumed to have reciprocal connections with every other, only connections from connotative size to connotative brightness are illustrated. Finally, but importantly, response selection is assumed to be based on connotative brightness.

Correct classification of the visual stimulus is assumed to develop over time. On a congruent trial, such as where the classification of a visual stimulus as bright requires the smaller key to be pressed, this key becomes progressively the more salient of the two. The increasing salience of this key's relative haptic size then translates into evidence for connotative smallness, which, in turn, translates into connotative brightness. In this way, the correct classification of the visual stimulus as being bright is reinforced, and selection of the correct response is facilitated. Conversely, on an incongruent trial, such as where the connotative brightness of the visual stimulus needs to be confirmed by pressing the bigger of the two keys, the increasing salience of the relative haptic size of this key provides evidence for connotative bigness, which then translates into evidence for connotative darkness. However, this evidence for connotative darkness contradicts the connotations of brightness originating with the visual stimulus, thereby impeding correct stimulus classification and creating conflicting response tendencies.

Further research

The proposal that response selection in the brightness classification task is based, at least in part, on the connotations of brightness and darkness indicated by the visual stimuli is amenable to experimental validation. So also is the proposal that the mapping of sensory-perceptual features onto dimensions of connotative meaning is context sensitive. Although in the present study, we provided evidence in support of this proposal in relation to visual brightness, the same also should apply to haptic size, despite this being a task irrelevant feature. The categorization of a response key as big or small will, to some extent at least, reflect its size relative to the key with which it is paired, rather than its absolute size. We have evidence for this from an additional study in which the wooden ball of intermediate size was used as one of

the response keys. It was observed to behave as a small key (in terms of how it contributed to a congruity effect) when paired with the biggest of the three wooden balls, but as a big key when paired with the smallest of the three balls.

Finally, with regard to the possibility that the size–brightness correspondence reflects a natural co-occurrence between the two features, the absence of evidence one way or the other means that contrasting claims can be made. In the present discussion, it is the lack of evidence that has been emphasized. However, a real need exists for a comprehensive programe of research to reveal the existence of natural co-occurrences, even those that are generally assumed to exist (e.g., between size and pitch), and to assess the consistency with which they support a particular correspondence. For example, although several natural co-occurrences might link size and brightness, they might do so in contradictory ways, some aligning small with bright, others aligning it with dark. In the context of a study in which a correspondence between surface brightness and weight was demonstrated, Walker, Francis, and Walker (2010) described a very tentative assessment of the existence of relevant co-occurrences. They reported having obtained samples of rocks (i.e., pebbles), and of different types of wood, and measuring both the surface brightness and density of each specimen. In the case of pebbles, they reported no association between surface brightness and weight, and a geologist colleague indicated that, in principle, an association would not be expected. With regard to different types of wood (taken from Edlin, 1969), Walker et al. reported a very modest, although significant, association between surface brightness and weight, with darker timber tending to be heavier than lighter colored timber. However, this association arose because of the exceptional heaviness and darkness of ebony. Furthermore, as Edlin pointed out, different types of wood do not vary in density when trees are freshly cut, only after they have been thoroughly air dried. Studies such as these, along with an examination of what, in principle, should be expected, have the potential to provide a sounder basis from which claims can be made regarding whether cross-sensory correspondences can, or cannot, have their origins in natural co-occurrences.

Appendix

Illuminant brightness and loudness are linear dimensions of magnitude reflecting the perceived strength of visual and auditory stimuli. In each case, it is clear which end of the dimension is least strong (i.e., no light, no sound) and how a stimulus should be changed to make it more or less strong. Accordingly, when people match different levels of illuminant brightness to different levels of loudness, and each of these to other magnitude dimensions—such as hand-grip force and size—very consistent correspondences emerge in which the dimensions arrange themselves with their more ends aligned (see Marks, 1978).

Although illuminant brightness and surface brightness are both commonly referred to as *brightness*, they are qualitatively distinct as dimensions of perceptual experience. This is reflected in those aspects of the English language used to refer to the properties of things and to feature dimensions (see Smith & Sera, 1992). In common with other magnitude dimensions, the *more end* of illuminant brightness is assigned a neutral, unmarked term (i.e., *bright*), which also serves as the term for the dimension as a whole (i.e., *brightness*), whereas the *less end* is assigned a marked term (i.e., *dim*). Thus, we talk about more or less brightness, but not more or less dimness. By contrast, the linguistic structure by which we talk about surface brightness is different. There is not an unmarked term for one end of the dimension, nor for the dimension as a whole (i.e., we can talk equally naturally about how light or how dark an object is), and we don't refer to different levels of surface brightness as being more or less gray (in the same way that we don't talk about a sound having more or less pitch, because pitch also is not a magnitude dimension).

Although both types of brightness follow a linear order (from dim to bright, and dark to light), it is only for illuminant brightness that this coincides with a gradation in magnitude or strength (i.e., only for illuminant brightness is brighter consistently more than dimmer). Thus, when Smith and Sera (1992) asked adults to indicate which of two contrasting values for surface brightness (presented either as samples of achromatic paper or the words *light* and *dark*) was *more than* the other, participants were equally likely to judge lighter or darker to be the *more end*. Similarly, when Marks (1974) and Smith and Sera asked people to match sounds varying in loudness to contrasting levels of surface brightness, the same inconsistency was apparent, with some people aligning lighter with louder (more than), and others aligning it with quieter (less than). And, with a different task situation, Marks (1982) confirmed that whereas illuminant brightness aligns itself consistently with loudness as a magnitude dimension, surface brightness does not. When he created metaphorical verbal expressions combining visual and auditory terms, he found that although words referring to contrasting levels of illuminant brightness (e.g., *dim* and *bright*) consistently altered the imagined loudness of a named sound with which it was combined (e.g., *dim sound of violin*), words referring to different levels of surface brightness (e.g., *dark*, *gray*, and *black*) did not (e.g. *black sound of violin*). Instead, as Marks (1982, p.188) observed:

"There is a very real sense in which black, though very dark, nevertheless is an intense color, just as white is. If both endpoints of the dark-bright continuum indicate intense colors, then this visual scale does not represent linearly increasing visual intensity."

Although surface brightness per se is not a magnitude dimension, it has sometimes been mistakenly seen as such because the level of contrast in surface brightness between a stimulus and its immediate background does appear to be a magnitude dimension. The significance of contrast in this regard became apparent to Marks when he observed participants in a visual–auditory matching task reverse the direction in which they aligned surface brightness with loudness when the background on which the paper samples of achromatic color were placed was switched from black to white (Marks, 1974). Confirmation was provided in another study in which loudness again served as the index of magnitude (Wicker, 1968). Wicker presented participants with paper samples of different colors on a medium gray surface, along with simple sounds varying in loudness and pitch. Participants were asked to indicate how similar each color was to each sound. Wicker found that whereas lighter colors were aligned with higher pitch sounds, they were not aligned with louder (*more*) or quieter (*less*) sounds. However, the contrast in lightness between a color sample and the immediate background against which it appeared did align itself with loudness, with more contrast being linked to more loudness. The association between brightness contrast and loudness reinforces the notion that black and white can both appear to be more than midgray, but not more or less than each other. Further evidence to this effect can be gleaned from two studies by R. Cohen Kadosh and colleagues, in which they explored the relationship between surface brightness and a more abstract index of magnitude—that is, the numerical value (size) of a visually presented digit. They too observed that whereas higher levels of brightness contrast between a digit and its background is associated with bigger numerical values (more than), the surface brightness of the digits themselves is not (R. Cohen Kadosh & Henik, 2006; Cohen Kadosh, Cohen Kadosh & Henik, 2008).

Marks (1987) confirmed that the qualitative differences in the behavior of illuminant and surface brightness extend to the computer presentation of visual stimuli, and to situations in which the stimuli interact with concurrent sounds to yield various types of congruity effect in speeded classification tasks. He presented contrasting levels of visual brightness in each of two ways. First, he presented small rectangles varying in brightness in the center of a computer screen that was otherwise dark (i.e., black) and that remained so throughout the experiment. Second, he presented similar rectangles on a midgray background and arranged for them to be lighter or darker than the background. Whereas Marks (1987) regarded the first manipulation as varying illuminant brightness (with the stimuli appearing as "luminous spots," p. 385), he regarded the second manipulation as varying surface brightness (with the stimuli appearing as "dark or light surfaces," p. 385). When these visual stimuli were used in speeded classification tasks, Marks (1987) confirmed qualitatively different patterns of congruity effect for the two types of brightness manipulation. Whereas illuminant brightness interacted with both pitch and loudness to yield the expected congruity effects (i.e., people responded more quickly and accurately when brighter visual stimuli appeared with higher pitched or louder sounds), surface brightness interacted only with pitch. The fact that the manipulation of surface brightness did not interact with loudness to yield a congruity effect was interpreted by Marks (1987) as further evidence that it does not represent linearly increasing intensity (i.e., it is not a magnitude dimension).

Because the present study was intended to explore congruity effects involving surface brightness, Marks' (1987) strategy for isolating it from illuminant brightness was adopted. That is, visual stimuli varying in brightness were presented, on a computer screen, on an otherwise midgray background. Three levels of brightness were darker than the background, and three levels were brighter than the background, and this is how they had to be classified.

To confirm that it is surface brightness, rather than illuminant brightness, that is being manipulated in this way, and that this is not a manipulation of stimulus magnitude, we asked a group of 32 undergraduate students, who were not involved in any other part of the present study, to say which of two stimuli from our speeded classification studies, appearing on the same midgray background, was "stronger or more than" the other. Specifically, they were asked to say which of the two darker stimuli that differed most in brightness (i.e., the darkest darker gray and the lightest darker gray), and which of the two brighter stimuli that differed most in brightness (i.e., the lightest lighter gray and the darkest lighter gray) was "stronger or more than" the other. The outcome was clear: Surface brightness, as operationalized by Marks (1987) and ourselves, is not a magnitude dimension. Instead, black and white are both stronger and more than midgray. Thus, twenty-nine participants judged the darkest of the darker grays to be more than the lighter of the darker grays, but the lightest of the lighter grays to be more than the darker of the lighter grays. The responses of two of the remaining participants indicated that lighter was more than darker, regardless of their relationship to the background. The responses of the final participant indicated that darker was more than lighter, regardless of their relationship to the background.

References

Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General, 138,* 351–367.

Cohen Kadosh, R., Cohen Kadosh, K., & Henik, A. (2008). When brightness counts: The neuronal correlate of numerical-luminance interference. *Cerebral Cortex, 18,* 337–343.

Cohen Kadosh, R., & Henik, A. (2006). A common representation for semantic and physical properties: A cognitive-anatomical approach. *Experimental Psychology, 53,* 87–94.

Cohen Kadosh, R., Henik, A., & Walsh, V. (2007). Small is bright and big is dark in synaesthesia. *Current Biology, 17,* R834–R835.

Collier, W. G., & Hubbard, T. L. (2001). Judgements of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicology, 17,* 36–55.

Edlin, H. L. (1969). *What wood is that? A manual of wood identification*. Ammanford, Carmarthenshire, Wales: Stobart Davies.

Eitan, E., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition, 114,* 405–422.

Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *American Journal of Psychology, 109,* 219–238.

Karwoski, T. F., Odbert, H. S. (1938). Color-music. *Psychological Monographs, 50* (Whole No. 222).

Karwoski, T. F., Odbert, H. S., & Osgood, C. E. (1942). Studies in synesthetic thinking: II. The role of form in visual responses to music. *Journal of General Psychology, 26,* 199–222.

Kohler, W. (1929). *Gestalt Psychology*. New York, NY: Liveright.

Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalences in early infancy: Auditory–visual intensity matching. *Developmental Psychology, 16,* 597–607.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *American Journal of Psychology, 87,* 173–188.

Marks, L. E. (1975). On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin, 82,* 303–331.

Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities*. New York, NY: Academic Press.

Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 177–193.

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 384–394.

Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 85–105). Cambridge, MA: MIT Press.

Martino, G., & Marks, L. E. (2001). Synaesthesia: Strong and weak. *Current Directions in Psychological Science, 10,* 61–65.

Meier, B. P., Robinson, M. D., & Caven, A. J. (2008). Why a big Mac is a good Mac: Associations between affect and size. *Basic and Applied Social Psychology, 30,* 46–55.

Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white. *Psychological Science, 15,* 82–87.

Meier, B. P., Robinson, M. D., Crawford, L. E., & Ahlvers, W. J. (2007). When "light" and "dark" thoughts become light and dark responses: Affect biases brightness judgments. *Emotion, 7,* 366–376.

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch–object correspondences in young children. *Cognitive, Affective, and Behavioral Neuroscience, 4,* 133–136.

Osgood, C. E. (1960). The cross-cultural generality of visual–verbal synesthetic tendencies. *Behavioral Science, 5,* 146–169.

Page, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association, 58,* 216–230.

Perrott, D. R., Musicant, A., & Schwethelm, B. (1980). The expanding-image effect: The concept of tonal volume revisited. *Journal of Auditory Research, 20,* 43–55.

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin, 132,* 416–442.

Robinson, E. J. (1954). The influence of photometric brightness on judgments of size. *American Journal of Psychology, 67,* 464–474.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology, 12,* 225–239.

Silvera, D. H., Josephs, R. A., & Giesler, R. B. (2002). Bigger is better: The influence of physical size on aesthetic preference judgments. *Journal of Behavioral Decision Making., 15,* 189–202.

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L., & Oakley, D. A. (2005). Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cognitive Neuropsychology, 22,* 1069–1085.

Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology, 24,* 99–142.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73,* 971–995.

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

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. London, England: John Wiley.

Von Hornbostel, E. M. (1931). Uber Geruchshelligkeit. *Pflugers Archiv fur die Gesamte Physiologie, 227,* 517–538.

Walker, P., Francis, B. J., & Walker, L. (2010). The brightnessweight illusion: Darker objects look heavier but feel lighter. *Experimental Psychology, 57*, 462–469.

Walker, P., & Smith, S. (1984). Stroop interference based on the synaesthetic qualities of auditory pitch. *Perception, 13,* 75–81.

Walker, P., & Smith, S. (1985). Stroop interference based on the multimodal correlates of haptic size and auditory pitch. *Perception, 14,* 729–736.

Walker, L., Walker, P. (2012). A common scheme for cross-sensory correspondences across stimulus domains. *Perception.*

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

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex, 42,* 264–280.

Wicker, F. W. (1968). Mapping the intersensory regions of perceptual space. *American Journal of Psychology, 81,* 178–188.

Zigler, M. J. (1930). Tone shapes: A novel type of synaesthesia. *Journal of General Psychology, 3,* 227–287.