The Size-Brightness Correspondence: Evidence for Crosstalk among Aligned Conceptual Feature Dimensions

Peter Walker[1], Laura Walker[1] & Brian Francis[2]

[1] Department of Psychology, Lancaster University, UK

[2] Department of Mathematics and Statistics, Lancaster University, UK

Author Note

Correspondence to:

Peter Walker, Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK

e-mail: p.walker@lancaster.ac.uk

tel: +44 (0) 1524 593163

fax: +44 (0) 1524 593744

Abstract

The same core set of cross-sensory correspondences connecting stimulus features across different sensory channels are observed regardless of the modality of the stimulus with which the correspondences are probed. This observation suggests that correspondences involve modality-independent representations of aligned conceptual feature dimensions, and predicts a size-brightness correspondence, in which smaller is aligned with brighter. This suggestion accommodates cross-sensory congruity effects where contrasting feature values are specified verbally rather than perceptually (e.g., where the words *WHITE* and *BLACK* interact with the classification of high and low pitch sounds). Experiment 1 brings these two issues together in assessing a conceptual basis for correspondences. The names of bright/white and dark/black substances were presented in a speeded brightness classification task in which the two alternative response keys differed in size. A size-brightness congruity effect was confirmed, with substance names classified more quickly when the relative size of the response key needing to be pressed was congruent with the brightness of the named substance (e.g., when *yoghurt* was classified as a bright substance by pressing the smaller of two keys). Experiment 2 assesses the proposed conceptual basis for this congruity effect by requiring the same named substances to be classified according to their edibility (with all of the bright/dark substances having been selected for their edibility/inedibility, respectively). The predicted absence of a size-brightness congruity effect, along with other aspects of the results, supports the proposed conceptual basis for correspondences and speaks against accounts in which modality-specific perceptuomotor representations are entirely responsible for correspondence-induced congruity effects.

The Size-Brightness Correspondence: Evidence for Crosstalk among Aligned

Conceptual Feature Dimensions

Each of our senses is 'blind' to some features of objects and events. For example, vision does not capture the sounds objects make, olfaction tells us nothing about their weight, and audition cannot discern their colour. As objects and events are not always available to all of the senses at once, there is considerable interest in how such 'blind spots' are filled-in, correctly or incorrectly, when the modality best placed to provide the missing information is unable to do so. How is it that in everyday life we readily refer to, for example, the brightness of a sound, the loudness of a shirt, and the thickness and heaviness of a perfume, when audition, vision, and olfaction are, respectively, 'blind' to these features? It seems that stimuli encoded in different sensory channels can share some of their perceptual features: Sounds can share their brightness and loudness with visual stimuli; and, odours can share their thickness and weight with objects seen and felt.

There appears to exist a core set of systematic associations (*cross-sensory correspondences*) connecting stimulus features encoded in different sensory modalities (L. Walker, P. Walker, & Francis, 2012; P. Walker & L. Walker, 2012a; P. Walker, 2012a, b). These cross-sensory correspondences offer a potential basis for the filling-in of information missing from different sensory channels. For example, because auditory pitch and visual brightness enjoy a corresponding relationship, high-pitched sounds normally 'feel' as though they are emanating from bright objects, even if the source of the auditory information cannot be seen.

Brightness, thickness, sharpness, and weight are all feature dimensions and evidence indicates that it is the relative positioning of stimuli on such dimensions that is shared by stimuli encoded in different sensory channels (e.g., their relative rather than absolute brightness). For example, when people judge what cross-sensory features are possessed by

sounds of different pitch, relatively high-pitched sounds are deemed to be more active, brighter, faster, higher in space, lighter in weight, sharper, and smaller than are lower-pitched sounds (Collier & Hubbard, 2001; Eitan & Timmers, 2010; Hubbard, 1996; Marks, 1974; 1975; 1978; Mondloch & Maurer, 2004; Perrott, Musicant, & Schwethelm, 1980; P. Walker & Smith, 1984). And when they draw music they are listening to, they draw lines and forms that are higher on the page, thinner, brighter, smaller, and more angular (sharper) the higher in pitch and/or faster in tempo is the music (Karwoski, Odbert, & Osgood, 1942; Kussner & Leech-Wilkinson, 2013).

The particular clustering of cross-sensory features linked to contrasting levels of auditory pitch is also observed when other types of stimulus contrast are explored, such as angular vs. curved visual shapes (P. Walker, 2012a; L. Walker, P. Walker, & Francis, 2012), bright vs. dark visual objects (P. Walker, 2012b; P. Walker, Francis, & L. Walker, 2010; L. Walker, P. Walker, & Francis, 2012), and small vs. large objects explored by touch alone (P. Walker & Smith, 1985; L. Walker, P. Walker, & Francis, 2012). The consistent appearance of the same clustering suggests that the feature dimensions involved in cross-sensory correspondences are aligned with each other in the same way whatever stimulus contrast is being explored (L. Walker, P. Walker, & Francis, 2012). This in turn accords with the view that the dimensions are modality-independent (i.e., amodal) (see the Discussion of Experiment 2 below), and that they are therefore well placed to provide a basis for the same cross-sensory features to be shared by stimuli encoded in different sensory channels.

Karwoski, Odbert, and Osgood (1942) propose that elementary stimulus features (e.g., visual surface brightness, visual angularity, auditory pitch) are rich in conceptual (semantic) connotations, and that the dimensions along which the features lie are aligned with each other

in ways that define the correspondences evident in cross-sensory induced imagery.[1] With regard to how the alignment of these conceptual dimensions shape such imagery, Karwoski et al. propose 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
> (*op. cit., p.* 217).

In this way, Karwoski et al. anticipate claims that cross-sensory correspondences can involve the conceptual representation of elementary stimulus features (see Martino & Marks, 1999; Melara & Marks, 1990; P. Walker & Smith, 1984). They also anticipate recent claims that such correspondences involve crosstalk (cross-activation) between correspondingly positioned feature values on different conceptual dimensions, and that the dimensions involved include those evident in the correspondences emerging when contrasting levels of auditory pitch are explored (see P. Walker, 2012a,b; P. Walker & L. Walker, 2012; L. Walker, P. Walker, & Francis, 2012) (see Figure 1).

---

[1] The dimensions of core features being considered here are not the three major factors of evaluation, potency and activity that were later promoted by Osgood as the basis of universals in meaning (with *universal* referring to generality across all types of stimuli and across all cultures) (see Osgood, Suci, & Tannenbaum, 1957) (see the Discussion to Experiment 1 below).

ALIGNED CONCEPTUAL FEATURE DIMENSIONS
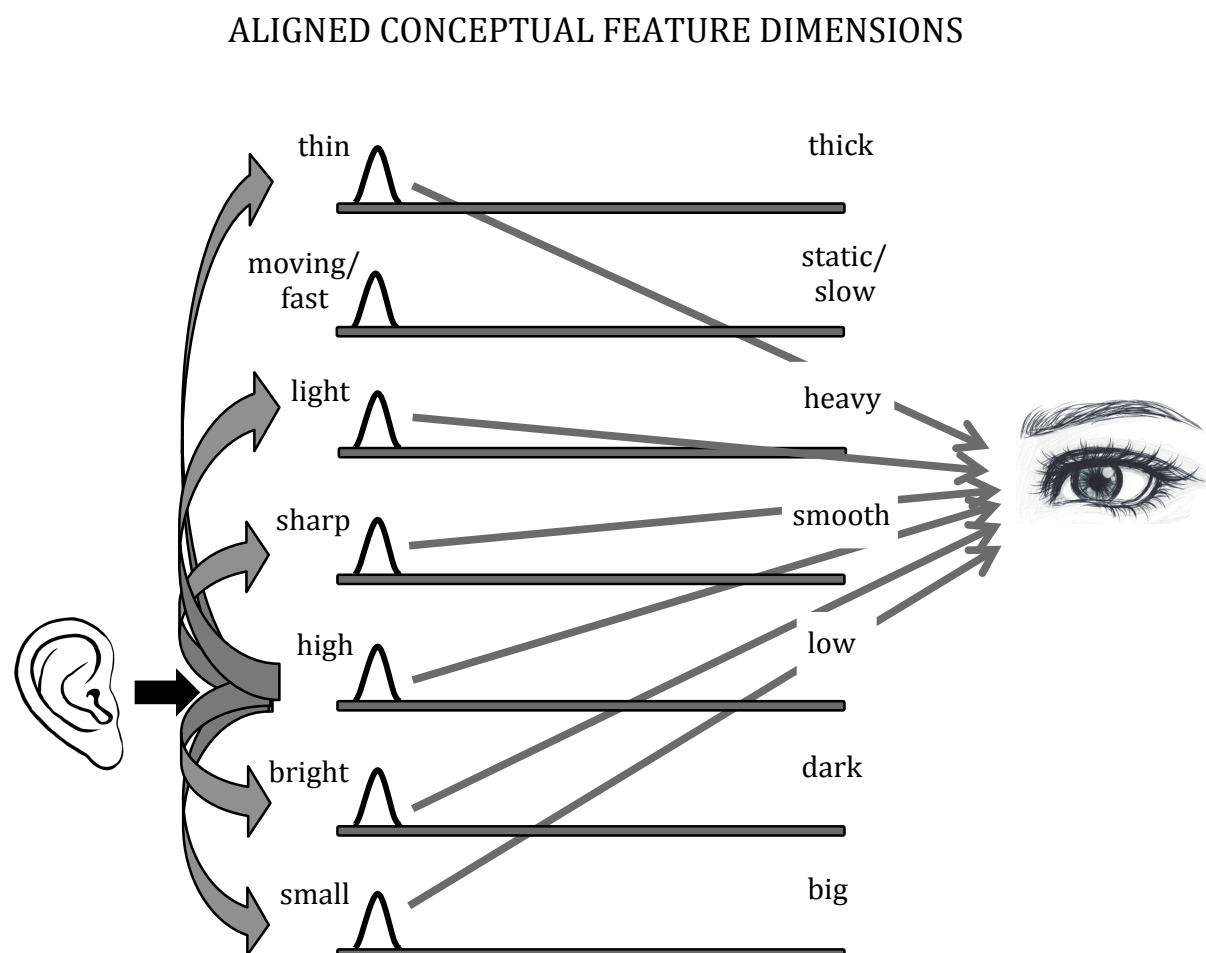
Figure 1 Correspondences evident in the visual imagery induced by sounds of contrasting pitch are thought to arise from the alignment, en bloc, of several conceptual dimensions (based on Karwoski, Odbert, & Osgood, 1942). Here, a relatively high-pitched sound induces visual images that, amongst other things, are relatively high in space, thin, sharp, bright and small. Though not shown here, it is assumed that extensive bi-directional activation occurs across corresponding places on all the dimensions. For example, when objects contrasting in size are explored haptically, not only are contrasting values of size activated, but corresponding features on other dimensions also are activated (e.g., the smaller object will induce other cross-sensory features, including *high*, *bright*, and *moving*).

Notwithstanding these recent claims regarding a key role for conceptual representations in cross-sensory correspondences, their involvement remains to be confirmed. In his tutorial review of cross-sensory correspondences, for example, Spence (2011) elects to highlight three non-semantic bases for correspondences, whilst at the same time acknowledging that correspondences might sometimes be rooted in the semantic representation of basic stimulus features. As evidence for the latter, Spence points to demonstrations of cross-sensory correspondences, typically using speeded classification tasks, in which at least some elementary stimulus features are presented verbally (e.g., with the words *high* and *low* replacing high and low pitched tones) (see also Gallace & Spence, 2006; Martino & Marks, 1999; Melara & Marks, 1990; P. Walker, 2012a; P. Walker & Smith; 1984; 1985).[2] It is this kind of evidence that prompted Martino and Marks (1999) to

---

[2] Sound symbolism in language relies on the potential for feature values presented verbally to engage with cross-sensory correspondences (see Marks, 1978, 1982). This is the case, for example, when the size of a table is communicated verbally using a relatively high-pitched (small) or low-pitched (big) vowel sound, thereby promoting *mil* and *mal* as appropriate names for a small table and big table, respectively (Sapir, 1929). In several early studies of sound symbolism it was shown that people could guess the meanings of antonyms from unfamiliar languages. Kunihira (1971), for example, demonstrated how English language college students, when presented with the Japanese equivalent of the antonym pair *bright-dark,* were able to guess significantly well which was the Japanese word for bright, and which was the Japanese word for dark. Brown and Nuttall (1959) did likewise for English language participants in relation to Chinese and Hindi antonyms. Significantly, however, their participants were unable to link Chinese and Hindi antonyms directly. Brown and Nuttall explain this limitation as being a consequence of their participants not having available the meaning of one of the pairs of antonyms. This implies, of course, that the

claim that cross-modal interactions can arise after information from different modalities is

recoded into an abstract format common to perceptual and linguistic systems, a format they

labelled *semantic* (*op. cit.,* p. 64).

Speeded classification tasks are an important context in which cross-sensory

correspondences are observed to impact on behaviour (see Marks, 2004). When people

classify stimuli on the basis of a criterial feature (e.g., classify a visual stimulus according to

whether it is bright or dark), they are influenced by whether an accompanying incidental

stimulus has corresponding (congruent) or non-corresponding (incongruent) features (e.g.,

whether an accompanying sound is high in pitch or low in pitch). More specifically, people

respond more quickly and accurately when the criterial and incidental feature values are

congruent (in correspondence) with each other, rather than when they are incongruent. For

example, people are faster to classify a visual stimulus as bright when it is accompanied by a

high-pitched sound (a bright sound), rather than a low-pitched sound (a dark sound) (Marks,

1987).

P. Walker and Smith (1984, 1985), and later Melara and Marks (1990) and Gallace

and Spence (2006), explore correspondence-induced congruity interactions in situations

where one of the interacting features is specified verbally (e.g., the words *HI* and *LO* are

presented either as printed text or as speech) and one non-verbally (e.g., the spatial elevation

of the word on the computer screen is high or low, or the overall auditory pitch of the spoken

word is high or low). Melara and Marks in particular seek to determine the type of

representation (e.g., visuo-spatial, auditory, graphemic, phonetic, lexical,

semantic/conceptual) on which correspondences can be based, looking specifically for

---

ability to exploit sound symbolism to identify equivalent antonyms from different languages

can be mediated by semantic coding, rather than only by directly comparing the perceptual

features of the words.

evidence that this can be semantic. On several counts they propose that this was the case in their study. First, the congruity effects they observed were symmetric, occurring whether the criterial feature was the feature specified verbally or the feature specified non-verbally. Second, the congruity effects were contingent on at least two contrasting feature values on each dimension being presented within a block of trials (which is taken as confirmation that it is the relative positioning of each feature value on its dimension that is critical for the correspondence, rather than its absolute values).[3] Third, though Melara and Marks acknowledge the potential for some low-level features of words (e.g., the pitch of their vowel sounds or the angularity of their letter forms) to induce cross-sensory interactions, they point out that in their experiments the words *HI* and *LO* interacted in the same way with different non-verbally presented features regardless of the sensory channel through which the latter were encoded. They considered it unlikely that sensory-perceptual levels of representation could have been involved in all the correspondence-based interactions they observed, leaving semantic representations as the only viable option.

 P. Walker (2012a) also observed correspondence-induced congruity effects with a mix of verbally and non-verbally presented feature values. He presented to-be-classified words inside novel outline shapes that were either angular or curved. The words referred to contrasting levels of auditory pitch, brightness, or hardness, and it was on the basis of each of these contrasts that participants classified the words. The congruity effects P. Walker observed reflected underlying interactions between the concept of sharpness (realised through the varying angularity of the shape), and the concepts of elevation, brightness, and hardness.

---

[3] Melara and Marks presume that the cross-sensory mapping of features underlying congruity effects involving sensory-perceptual representations will be based on the absolute values of features, and so will not require other feature values to provide a context for their relative positioning on a dimension.

Specifically, the angularity of the outline geometric shape within which a to-be-classified word appeared interacted with the conceptual connotations of the word to yield a correspondence-induced congruity effect. For example, his participants found it easier to classify a word as referring to a high-pitched (sharp) sound when it appeared within an angular (sharp) shape.

Conceptual Coding and the Size-Brightness Correspondence

The particular clustering of cross-sensory features reappearing across a range of situations, together with the assumed transitivity of cross-sensory correspondences (see Marks', 1978, account of Hornbostel's, 1931, demonstration of the transitivity of the correspondences between surface brightness, odour, and auditory pitch), predicts a correspondence between size and brightness, with smaller aligning with brighter. P. Walker and L. Walker (2012) explain why this correspondence is unlikely to have a non-conceptual basis, and provide evidence for its induction of a congruity effect in a brightness classification task. Their participants were presented with individual circles at one of six levels of brightness on a mid-grey background. Three levels were brighter than the background, and three were darker than the background, and participants had to classify each circle as quickly as possible according to whether it was brighter or darker than the background. They confirmed their decision by pressing one of two hidden response keys with their left or right hand. As a task irrelevant aspect of the situation, the response keys differed in size, so that on any trial the key needing to be pressed was either the smaller or larger of the two keys. P. Walker and L. Walker observed the congruity effect predicted from the correspondence between size and brightness, with participants classifying brighter (darker) circles more quickly when the key needing to be pressed was the smaller (bigger) of the two.

Despite the contrasting levels of brightness and size being presented non-verbally, P. Walker and L. Walker (2012) reasoned that the congruity effect resulted from processes taking effect after the brightness classification of each circle (i.e., they were post-categorical). First, when a test circle was brighter or darker than the mid-grey 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 linked to 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 each stimulus. Second, because participants grasped the two response keys continuously throughout a block of trials, one of the two keys, and its relative size, became relevant as the key needing to be pressed only after a 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 likely to be the key needing to be pressed and congruity with one particular key size was not yet an issue. On these two counts, therefore, conceptual (post-categorical) coding appeared to mediate the correspondence-induced congruity effect.

Additional support for conceptual coding in the size-brightness congruity effect has since been reported by L. Walker and P. Walker (2015), who provide more direct evidence that it is the relative brightness of the circles, and the relative sizes of the response keys, that are implicated in the effect rather than the absolute values of these features. L. Walker and P. Walker show how the same circle can interact with key size as either a bright circle or a dark

circle depending on the brightness of the other visual stimuli with which it appears. Similarly, the same response key can interact with brightness either as a small key or as a big key depending on the size of the other key made available to participants. Observing the functional significance of context-sensitive, relative coding resonates with a recent study of the correspondence between auditory pitch and visuo-spatial elevation in which Chiou and Rich (2012) show that correspondence-induced congruity effects between non-verbal stimuli can reflect the relative coding of their feature values. As noted already (see Footnote 3), cross-sensory mappings based on the relative coding of feature values is generally taken as evidence of a conceptual basis for the mappings, whereas cross-sensory mappings based on absolute feature values is thought to indicate the involvement of sensory-perceptual representations (so that their impact does not require other feature values to provide a context for their relative positioning on their dimension) (e.g., Chiou & Rich, 2012; Gallace & Spence, 2006; Lunghi & Alais, 2013; Lunghi, Binda, & Morrone, 2010; Marks, 1987; Marks, Szczesiul, & Ohlott, 1986; Melara & Marks, 1990; Orchard-Mills, Alais, & van der Burg, 2013; Orchard-Mills, van der Burg, & Alais, 2013).

Focussing on the size-brightness correspondence, the present study was designed to confirm that, in principle, cross-sensory correspondences can have a basis in the interactions among conceptual representations of elementary stimulus feature dimensions. In Experiment 1, a version of P. Walker and L. Walker's (2012) brightness classification task was adopted, but with contrasting levels of brightness specified verbally. This was achieved by presenting words referring to either bright (white) substances, or dark (black) substances, and asking participants to classify them according to the brightness of their referent. Participants communicated their classification decision by pressing one of two response keys that,

incidentally, differed in size.[4] The idea that cross-sensory correspondences reflect

interactions among representations having an abstract format common to perceptual and

linguistic systems (i.e., Martino & Marks', 1999, *semantic coding account*), through the

property of transitivity this entails, predicts a size-brightness congruity effect, with the

relative size of the response key needing to be pressed on a particular trial interacting with the

level of brightness associated with the word's referent. In particular, participants should

classify words relatively more easily (quickly and accurately) when a word referring to a

bright (dark) substance requires the smaller (bigger) of the two keys to be pressed.

In Experiment 2, the same two sets of words were classified, but now on a different

basis making no reference to brightness, or indeed to any of the core cross-sensory feature

dimensions involved in correspondences. Thus, the names of bright and dark substances used

in Experiment 1 were now classified according to the edibility of their referents. Because all

of the bright substances had been chosen for their edibility, and all of the dark substances for

their inedibility, everything from Experiment 1 bar the basis for classification was kept in

place for Experiment 2. To the extent that correspondences, and the congruity effects they

induce, reflect interactions among conceptual representations of core cross-sensory feature

dimensions, rather than among conceptual representations more generally, a size-brightness

---

[4] Although studies of cross-sensory correspondences typically involve just two response

alternatives mapping onto two categories as the basis for stimulus classification (here bright

vs. dark), there is nothing in the theoretical framework being promoted here that restricts it to

binary distinctions. In principle, the framework extends to the classification of stimuli into

any number of categories (i.e., any number of distinct ordered categories into which a

dimension can be partitioned), whether this be the nine categories created by Dolscheid,

Shayan, Majid, and Casasanto, (2013), or the three and five categories created by Thompson

and Estes (2011) in different of their experiments.

congruity effect would not be expected with this alternative basis for classifying the words.

In addition, however, taken together, Experiments 1 and 2 also examine a very different way

of explaining cross-sensory correspondences and the congruity effects they induce. This

alternative explanation, derived from notions of embodied cognition, assumes that

interactions involving modality-specific perceptuomotor simulations are able to explain the

same evidence adequately, even when this evidence comes from situations in which the

targeted stimulus feature values are presented as words whose referents typically have these

values (see the Introduction and Discussion to Experiment 2 for a fuller account).

Experiment 1: Size-Brightness Correspondence with Brightness Presented Verbally

To confirm that conceptual representations of size and brightness can provide a basis

for their correspondence, and for the congruity effects this correspondence induces,

participants in Experiment 1 were presented with contrasting levels of brightness as the

names of substances that are typically either bright (white, or close to white) (e.g., *flour*) or

dark (black, or close to black) (e.g., *coal*) in colour.[5] They were asked to classify the names

as quickly and as accurately as possible according to the brightness of their referent by

pressing either the left or right of two response keys which, incidentally, differed in size. If

the size-brightness correspondence is, at least in part, based on conceptual representations

that can be accessed through either verbal or non-verbal stimuli, then participants should

respond more quickly and accurately when the size of the key needing to be pressed is

congruent with the brightness of the named substance (e.g., *sugar*-small, *soot*-big), than when

---

[5] Because the substance names do not double up as verbal labels for contrasting levels of size

(i.e., corresponding feature values do not link to the same lexical entities), a size-brightness

congruity effect could not have a lexical basis (cf. Spence, 2011) (see P. Walker, 2012a, for a

fuller exploration of this issue).

it is incongruent with it (e.g., *soot*-small, *sugar*-big). In other words, a congruity effect

equivalent to that previously observed with the non-verbal presentation of contrasting levels

of brightness (P. Walker & L. Walker, 2012) should be observed, consistent with the notion

that cross-sensory correspondences can involve representations having an abstract format

common to both perceptual and linguistic stimuli, rather than being exclusively sensory-

perceptual in form (Martino & Marks, 1999).

Method

Participants

Twenty-five Lancaster University students (20 females and five males) aged between

18 and 42 (mean age = 20.4 years) volunteered to participate in the experiment in exchange

for course credit or payment. All but two of the participants were right-handed by self-report.

Task, Materials, and Apparatus

Participants completed 240 trials, in each of which they were required to decide

whether the physical referent of a visually presented word was bright or dark. The visual

stimuli consisted of 10 brightness-related words obtained from the British National Corpus

website (University of Oxford, 2010). Five of the words referred to items (substances)

typically with a relatively high level of surface brightness (i.e., bright), and five referred to

substances typically with a relatively low level of surface brightness (i.e., dark). With the

average of their written and spoken word frequencies (per 100 million words) in parentheses

(taken from the British National Corpus website), the five bright words were *milk* (4737),

sugar (3694), *salt* (2945), *flour* (1036), and *yoghurt* (287), and the five dark words were *coal*

(5061), *soil* (4129), *ink* (793), *tarmac* (392), and *soot* (185). Though the average frequency

was comparable for the two sets of words, 2540 and 2112, for the bright and dark words,

respectively, word frequency was later entered as a fixed effects factor in linear mixed effects analysis of the results.

Few words are able to serve the purpose required of them, especially when their additional role in Experiment 2 is anticipated (see below). There was, therefore, little opportunity to ensure that the sets of bright and dark words were matched on all features, aside from frequency, having the potential to interact with response key size and, therefore, the potential to provide an alternative explanation for what would otherwise appear to be a size-brightness congruity effect. Rather fortuitously, however, the sets of bright and dark words that were available happened to be closely matched on: the average position of their vowels (i.e., the front-back aspect according to the International Phonetic Alphabet (IPA)); the average height of their vowels (i.e., the open-close aspect according to the IPA); and the visual angularity of their letter forms. Appendix A provides details of these and other features of the two sets of words, including how some of the feature values were assessed. With regard to the percentage of a word's consonants that were plosive in nature (a proxy for the abruptness (angularity) of their acoustic amplitude envelope, see Rhodes, 1994), there was a noticeable difference in the mean values for the two sets of words. However, the direction of this difference contradicts the predicted interaction between brightness and response key size (i.e., it is the darker words that are phonologically more angular, on which basis they would be congruent with the smaller key, rather than the bigger key). In addition, with regard to the visual size of the words, indexed by the number of letters they comprise, there is a 20% difference across the two sets of words. Because of the obvious potential link between word size and key size, word length also was entered as a fixed effects factor in the linear mixed effects analysis of the results to determine if it interacted with key size. Finally, with regard to the typical portion sizes in which the named substances are encountered, the direction of the difference across the bright and dark words conflates with the brightness

difference (i.e., the brighter substances are typically encountered in smaller portions than the darker substances, which in its own right could interact with response key size to yield what could be mistaken for a size-brightness congruity effect). With regard to this potential confound with the manipulation of brightness, typical portion size also was entered as a fixed effect in the linear mixed effects analysis of the results to determine if it interacted with key size.

The words were presented individually at the centre of a 20" computer screen (Apple PowerMac G5, Dual 2GHz), running version 2.1.1 of the PsyScript experiment generator programme. Each word appeared in uppercase and was displayed in black on a white background in a 50-point, Calibri font. Participants immediately classified the physical referent of each word as either bright or dark by pressing either the left or right of two keys that differed in size. The difference in the sizes of the two response keys was incidental to the speeded brightness classification task and was never mentioned by the experimenter.

The response keys comprised two smoothed wooden balls mounted onto micro-switches. The small ball had a diameter of 2.5 cm and the big ball had a diameter of 7.5 cm. The physical resistance of the two switches was adjusted until the authors judged that equal force was needed to close them. This required that a higher level of resistance was set for the big key (1000 gm) than for the small key (250 gm). The small key was also raised 3.75 cm from the table by a wooden block to ensure that the two balls were perceived (haptically) by the participants to be of equal spatial height. The sound made by closing the micro-switches was very quiet, and appeared to be identical for the two keys. Nevertheless, it was masked with a single beep sound presented whenever a key was pressed (through two Creative SBS250 2.5 watt stereo speakers located at either side of the computer screen). A thick black cloth was also used to cover the response keys at all times during the experiment, as a result of which participants never saw them.

Design and Procedure

Participants classified the words according to the brightness of their referent (e.g., *flour* = bright, *coal* = dark) and were informed that they would complete 240 trials, in each of which a word would be presented at the centre of the computer screen. They were told that their task was to decide as quickly and as accurately as possible whether the item to which the word refers is bright or dark. Half of the participants (12 right-handers and one left-hander) were required to press the left-hand key if the item was bright and the right-hand key if the item was dark, and half (11 right-handers and one left-hander) were assigned to the opposite brightness-hand mapping (i.e., right-hand key for bright and left-hand key for dark).

Participants completed four blocks of trials, and were given a 1-minute break between blocks. In a block of 60 trials, each of the 10 words appeared six times, in a randomly determined order that was generated afresh for each participant. Each word remained visible until participants made their brightness decision, and was followed by a blank interval of 1.5 seconds before the next word was presented. Participants did not receive feedback about the speed or accuracy 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 key size-brightness mapping. Thus, across the four experimental blocks, participants alternately pressed the small key for *bright* and the big key for *dark* (congruent mapping), or the small key for *dark* and the big key for *bright* (incongruent mapping). Which of these two mappings (small vs. big key for *bright*) was assigned to the first block of trials was counterbalanced across participants.

Results

The data were the accuracy and speed of participants' responses to the words. Accuracy levels and mean observed correct response times (RTs) (after replacing excessively long RTs with cut-off values set at 2.5 *SD*s above a participant's mean RT) obtained for the bright and dark words, calculated separately for the small and big key, are shown in Table 1.

Response Accuracy

The overall mean level of response accuracy was 98.6% (*SD* = 4.0%). The mean percentage of correct responses was not significantly higher for congruent trials (98.7%) than for incongruent trials (98.5%), Wilcoxon Signed Ranks Test *z* = 0.78, *p* = .22, one-tailed.

Table 1 Mean observed correct RT (*SE* in parentheses) and *p*(correct) for each of the *bright* and *dark* words in Experiment 1 according to whether it required a response on a congruently sized key or an incongruently sized key

Test word	Congruent	Incongruent	RT difference (inc-con)
bright			
milk	681 (18) 0.98	**673 (16) 0.99**	-8
sugar	680 (20) 0.99	**684 (26) 1.00**	4
salt	708 (19) 0.96	**724 (27) 0.97**	16
flour	713 (20) 0.99	**665 (17) 0.99**	-48
yoghurt	669 (16) 0.99	**668 (20) 0.99**	-1
dark			
coal	**667 (17) 1.00**	680 (15) 0.99	13
soil	**679 (20) 0.99**	712 (24) 0.97	33
ink	**669 (15) 0.98**	702 (19) 0.97	33
tarmac	**632 (14) 1.00**	685 (19) 1.00	53
soot	**663 (13) 0.98**	702 (19) 0.98	39

Note. Normal and bold text relate to trials on which the small key
and the big key needed to be pressed, respectively

Response Speed

Prior to statistical analysis, individual RTs were subject to reciprocal transformation (i.e., converted to response speed) to improve the normality of the residuals. R (R Core Team, 2012) and the package lme4 (Bates, Maechler & Bolker, 2014) were used to perform crossed linear mixed effects analysis of the relationship between response speed and the congruence between key size and brightness. Crossed mixed effects, or multi-level models, first proposed by Goldstein (1994) in an educational research setting, have been popularized for the analysis of psychology experiments by Baayen, Davidson and Bates (2008).

TRIAL (1 – 240), KEY SIZE (big, small), BRIGHTNESS (whether the named item was bright or dark), AGE, GENDER, PORTION SIZE (the typical portion size for each named item on a scale from 1 (very small) to 6 (very big)), WORD FREQUENCY, and WORD LENGTH (number of letters), were included as fixed effects factors. The two design factors that were counterbalanced across participants also were included, that is, to which hand (left or right) bright decisions were assigned in the first block of trials (HAND FIRST), and the size of key (big or small) assigned to bright decisions in the first block of trials (SIZE FIRST). The intercepts for PARTICIPANTS and WORDS were treated as having random effects on response speed. For all the analyses, visual inspection of Q-Q plots of the residuals did not reveal any obvious departures from normality.

Likelihood ratio tests compared a basic model against models in which the interaction between key size and each of word length, portion size, and brightness were added in turn as fixed effects factors. This allowed the significance for response speed of each type of congruence with key size (i.e., word-length congruence, portion-size congruence, and

brightness congruence) to be assessed. Where word length and portion size interacted significantly with key size, these interactions were added to the basic model against which later models were compared.

There was neither a significant WORD LENGTH X KEY SIZE interaction, $\chi^2(1) =$ 0.24, p = .62, nor a significant PORTION SIZE X KEY SIZE interaction, $\chi^2(1) = 2.76$, p = .10. There was, however, a significant BRIGHTNESS X KEY SIZE interaction, $\chi^2(1) =$ 5.00, p = .03, the nature of which confirmed the congruence effect predicted from the size-brightness correspondence, with congruence raising response speed by 0.017 decisions/sec (SE = 0.008), reflecting a 13 ms reduction in observed RT, from 685 to 672 ms (see Figure 2). Inspection of the coefficients for the effects of the congruence between key size and brightness, by word, confirmed that it facilitated decision speed for every word except *flour* (coefficient range = -0.006 to 0.033 decisions/sec, mean coefficient = -0.017, SD = 0.011).

Response Speed (decisions/sec)
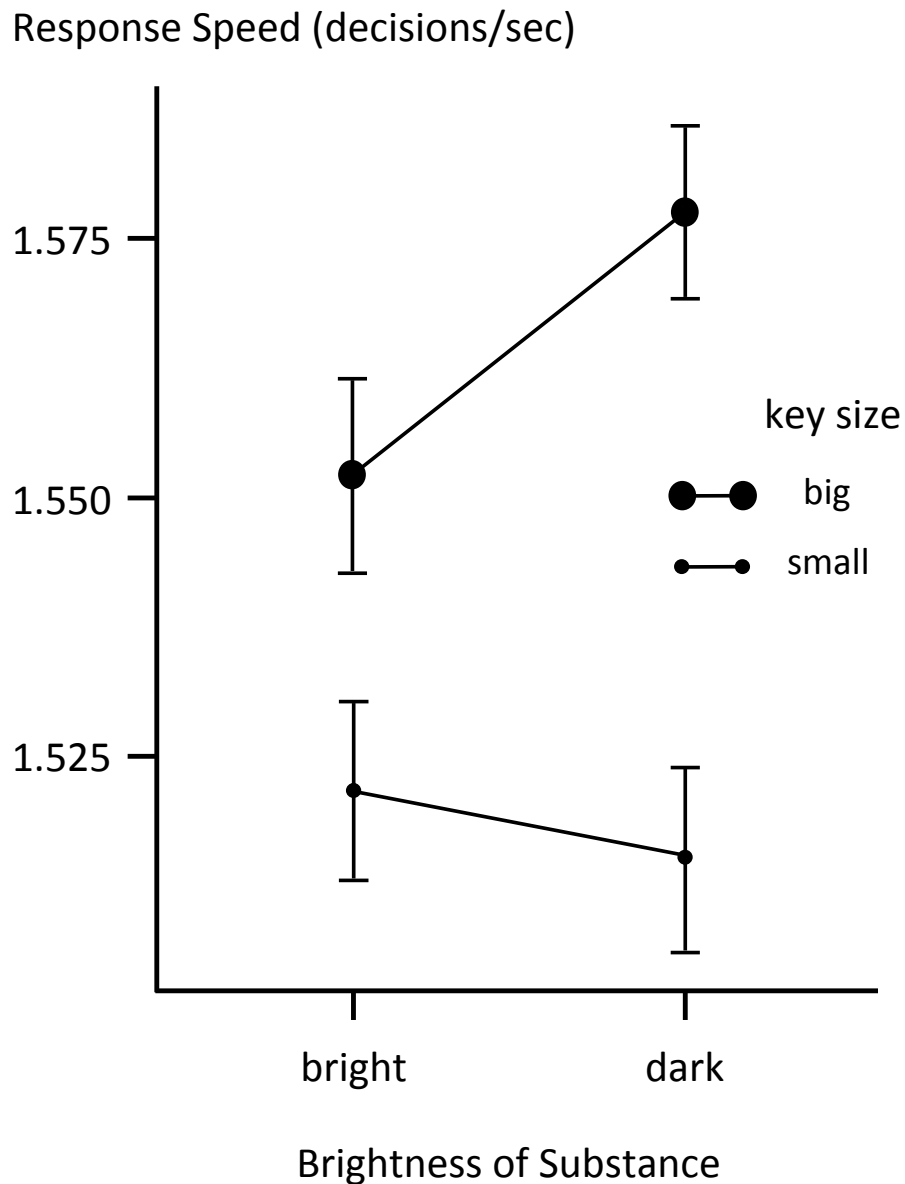
Figure 2 Experiment 1: Brightness classification. Mean observed response speed (decisions/sec) according to the level of brightness being classified and the size of the key on which the brightness classification decision is being communicated (bars indicate *SE*s).

Discussion

The results replicate the size-brightness congruity effect previously reported by P. Walker and L. Walker (2012), but now with contrasting levels of surface brightness indicated verbally as the names of substances that are typically bright (white) or dark (black) in colour.

Participants responded more quickly when the brightness of the named substance was congruent with the size of the key needing to be pressed (e.g., *sugar*-small, *soot*-big), rather than incongruent with the size of this key (e.g., *sugar*-big, *soot*-small). Because typical portion size was confounded with substance brightness in the set of substances being sampled (the brighter substances are typically encountered in smaller portions than are the darker substances), portion size was incorporated in the analyses as a fixed effects factor. It was confirmed that the size-brightness congruity effect remained significant when typical portion size was taken into account. Indeed, in the event, typical portion size did not interact with key size, and so could not provide an alternative explanation for what is being interpreted as a size-brightness congruity effect. The length (size) of the words themselves also did not interact with key size.

Observing a congruity effect equivalent to that observed when contrasting levels of brightness are presented non-verbally (P. Walker & L. Walker, 2012) supports the notion that the same conceptual representations of size and brightness are being accessed by non-verbal and verbal stimuli, and that these representations can underlie cross-sensory correspondences. More specifically, it appears that the observed size-brightness congruity effect reflects processes taking effect after the information from different sensory channels has been recoded into an abstract conceptual format common to perceptual and linguistic systems, rather than being based exclusively on sensory-perceptual representations (e.g., Martino & Marks, 1999). According to P. Walker et al.'s understanding of cross-sensory correspondences, this common format incorporates the representation of a core set of aligned feature dimensions and the mutual interactions (crosstalk) among them (P. Walker, 2012a; P. Walker & L. Walker, 2012; L. Walker, P. Walker, & Francis, 2012). Figure 3 illustrates how the size-brightness congruity effect observed in Experiment 1 is explained on this basis.
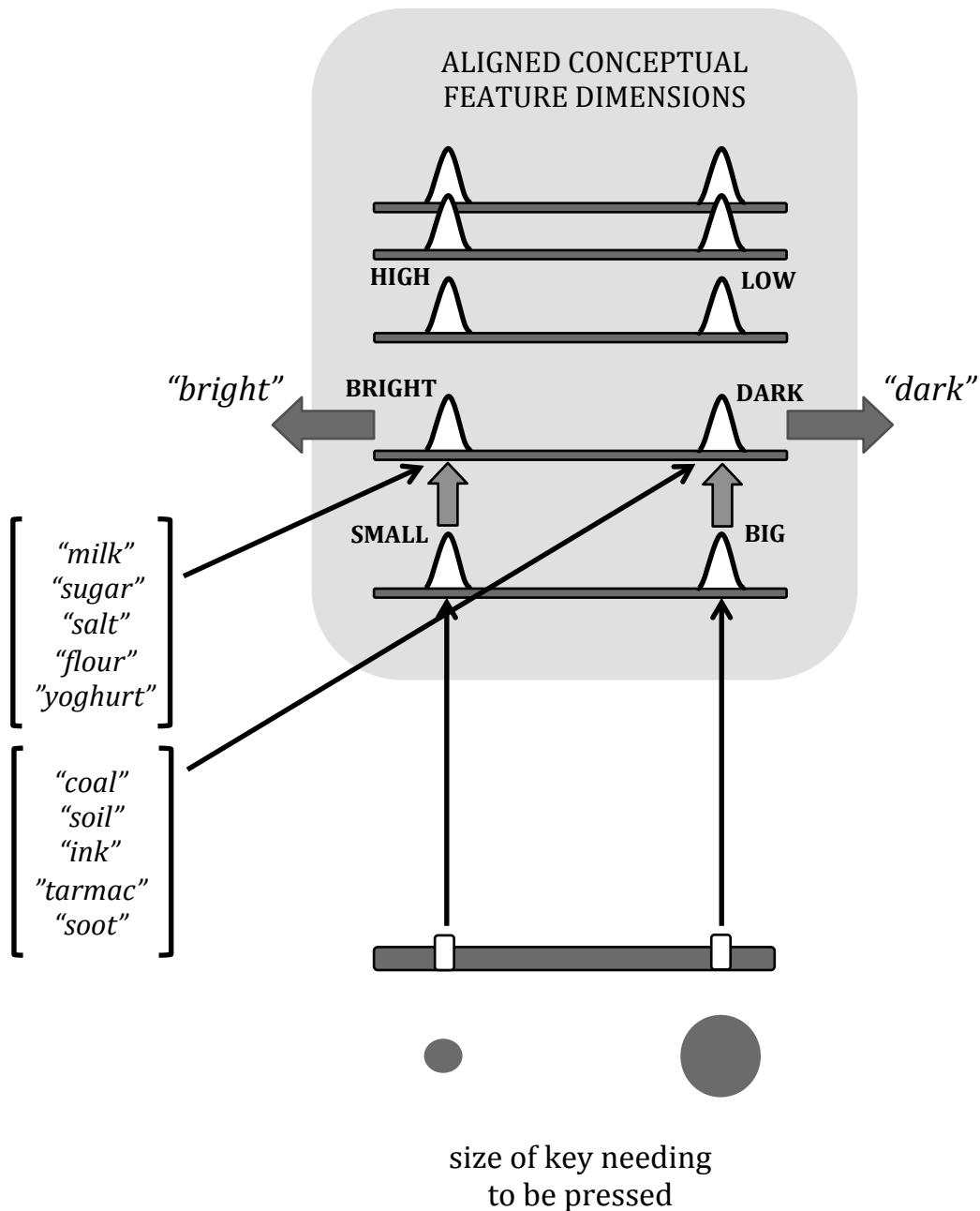
Figure 3 The functional components thought to be responsible for the size-brightness correspondence, and the congruity effect it induces during speeded brightness classification. Contrasting levels of visual brightness, presented verbally, map onto contrasting levels of conceptual brightness, while contrasting levels of haptic size map onto contrasting levels of conceptual size. It is assumed that the correct brightness classification of the word is based on the relative level of conceptual brightness emerging over time with the most supporting evidence. On a congruent trial, such as where the classification of a word as bright requires the smaller key to be pressed, the congruently sized key gradually becomes the

more relevant of the two keys, as it emerges that it is the key needing to be pressed. The increasing relevance of this key's relative size translates into evidence for conceptual smallness, which then translates, through cross-activation, into conceptual brightness. Classification of the word as referring to something bright is reinforced by this cross-activation, facilitating faster response selection. Conversely, on an incongruent trial, such as where the classification of a word as bright requires the bigger key to be pressed, the increased response relevance of the incongruently sized key as the key needing to be pressed feeds forward as evidence for conceptual bigness, which then translates into evidence for conceptual darkness. But this evidence for darkness contradicts the brightness developing from the word being classified, with conflicting response tendencies being induced and correct stimulus classification being slowed down.

P. Walker and L. Walker (2012) explain that the core set of aligned feature dimensions do not map on to (reduce to) the three dimensional factors emerging from Osgood's work on universals of meaning (i.e., the factors of *evaluation, potency,* and *activity*) (see Osgood, Suci, & Tannenbaum, 1957). They rule out evaluation on empirical grounds, and then point out that, as strength (magnitude) dimensions, potency and activity also are unlikely, not least because some of the dimensions involved in the core set of correspondences are not of this type. Most notable among these is auditory pitch (e.g., Smith & Sera, 1992). When Eitan and Timmers (2010) submitted the cross-sensory features associated with auditory pitch to principal components analysis, using the semantic differential technique, they confirmed evaluation, potency, and activity as three important underlying factors, but then observed them not to be the strongest predictors of the cross-sensory associations enjoyed by pitch. This status went to a factor, which included brightness, that Eitan and Timmers found difficult to conceptualize. Just how many core correspondences there are, and how they might mesh with a tripartite scheme such as that proposed by Osgood and his colleagues, therefore remains for further research to clarify.

Experiment 2: Switching the Basis for Stimulus Classification from

Brightness to Edibility

P. Walker et al.'s framework for understanding cross-sensory correspondences says

more than that they can be conceptually mediated (see P. Walker, 2012a; P. Walker & L.

Walker, 2012; L. Walker, P. Walker, & Francis, 2012). It also stipulates that it is the

crosstalk among representations of a core set of aligned feature dimensions that is at the heart

of cross-sensory correspondences. It follows from this that two conditions need to be in place

for cross-sensory correspondences to give rise to a congruity effect in speeded classification.

First, the classification decision needs to refer to one of the cross-sensory feature dimensions

involved in correspondences. Second, encoding of the incidental stimulus feature needs to

converge on the same feature dimension. As can be seen from Figure 3, both of these

conditions were in place in Experiment 1.

There are many conceptual bases on which named substances could be classified,

including, for example, whether they are expensive or inexpensive, manufactured or natural,

typically found in the home or not, and toxic or not. It follows from P. Walker et al.'s

framework that classifying substances on any such basis should not induce a correspondence-

based congruity effect because such classification would not contact any of the feature

dimensions involved in cross-sensory correspondences. This line of reasoning offers a way

of testing P. Walker et al.'s framework, and informed both the selection of bright and dark

substances for use in Experiment 1, and the strategy adopted in Experiment 2.

All of the bright substances named in Experiment 1 were selected because of their

edibility (i.e., they are foods, or ingredients added to food), and all of the dark substances

were selected because of their inedibility. This provides an alternative conceptual basis for

the classification of the words, while at the same time preserving their individual mappings

onto the two alternative responses. At the very least, therefore, Experiment 2 is able to

confirm that no word-level features inadequately controlled in Experiment 1 were responsible

for what is being interpreted here as a size-brightness congruity effect. If any such features

were responsible, then the same congruity effect should be observed in Experiment 2.

Experiment 2 goes further than this, however: Because the two pre-requisites for a

correspondence-based congruity effect are not satisfied, such an effect should not be

observed, despite the conditions pertaining in Experiment 1 otherwise remaining in place (see
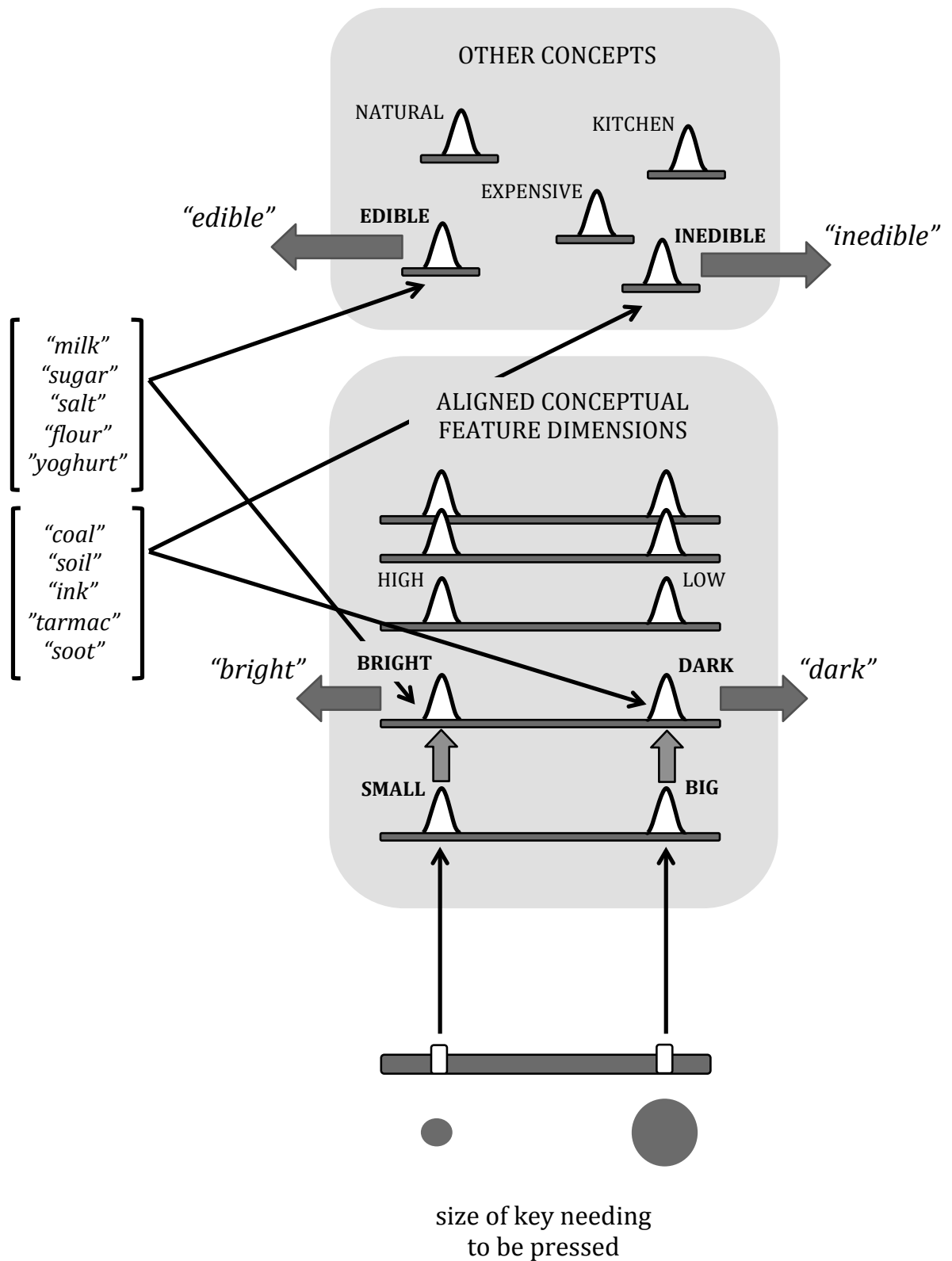
Figure 4).

Figure 4 Conceptual representations distinct from those concerned with aligned

conceptual feature dimensions can provide alternative bases for classifying the

same stimuli. Where they do, as envisaged here for edibility, classification decisions might remain uninfluenced by the information being registered concerning the cross-sensory features of concurrent stimuli, such as the relative size of the response key needing to be pressed. Where this is the case, a correspondence-induced congruity effect should not be observed.

Perceptuomotor simulations

As indicated in the Introduction, there is an alternative approach to explaining the evidence for cross-sensory correspondences, along with the congruity effects they induce, based entirely on modality-specific perceptuomotor representations. This approach starts by assuming that specifying a feature value verbally by naming an item for which the value is a typical attribute (e.g., salt is typically white) does not preclude sensory-perceptual representations from providing a basis for cross-sensory correspondences. Rather, such words are thought to activate sensory-perceptual representations that then are available to mediate correspondence-induced congruity effects. Experiment 2 contributes significantly to the assessment of this alternative approach.

This alternative approach follows from notions of embodied cognition and, more specifically, from the idea that the referent of a word can be represented as a reinstatement (simulation) of the perceptuomotor experiences induced during a previous encounter with the referent itself (e.g., Barsalou, 1999, 2009; Solomon & Barsalou, 2004). Where an item is being named, it is modality-specific perceptuomotor experiences linked to the item, or, more accurately, to an exemplar of the named item category, that is of most interest, rather than experiences linked to the word itself. Thus, seeing the word *salt* induces a re-experience of what was seen, heard, tasted, and felt during a previous encounter with an instance of this substance. Critically, where the values for different modality-specific features of named items correlate (correspond) with each other in the real world and, therefore, in our experiences of them, they will also correlate in perceptuomotor simulations. It is this

correlation that is assumed to underlie a cross-sensory correspondence. For example, for the

substances named in the present study, a correlation exists between visual surface brightness

and (visual and haptic) portion size, with brighter aligning with smaller. It is assumed that

this association will be reflected in the concurrent presence of aligned feature values in

perceptuomotor simulations of the substances (e.g., brighter visual re-experiences will tend to

co-occur with smaller (visual and haptic) re-experiences). This co-occurrence will have the

potential to prime responses when the relative size of the key needing to be pressed matches

the visual and/or haptic size associated with the level of brightness specified in the

perceptuomotor simulation. That is, where the level of brightness being re-experienced is

associated with a small (big) portion size, responses on the key matching the portion in size

will be primed (see Figure 5). In this way, the cross-sensory feature associations reflected in

perceptuomotor simulations could support a correspondence-based congruity effect, provided

they are combined with a mechanism able to link these feature values to representations of

other concurrent stimuli (here the two response keys). In this way there would be no need to

implicate modality-independent representations of named concepts, such as the abstract

forms of representation assumed by Martino and Marks (1999) to be shared by perceptual and

linguistic stimuli.
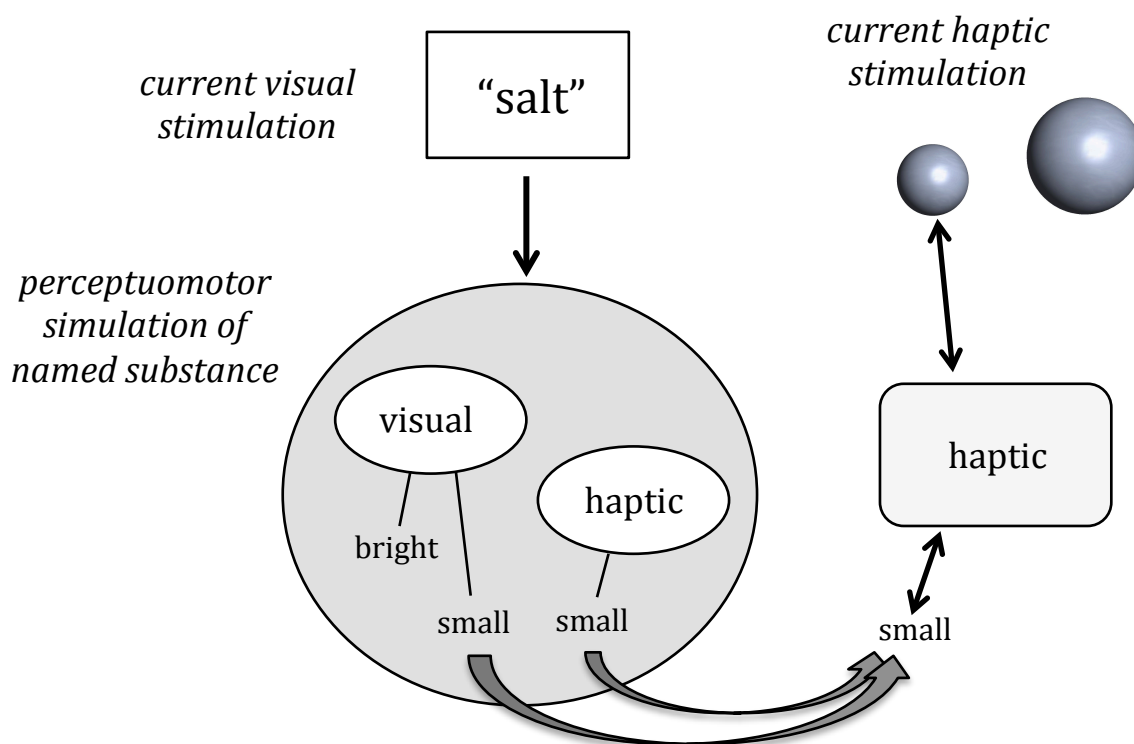
Figure 5 How co-occurrences among the features incorporated in a modality-specific perceptuomotor simulation of a named substance, here brightness and typical (visual and haptic) portion size, might come to activate responses linked to congruent values of haptic size. In this way, a congruity effect would be induced by a cross-sensory correspondence solely through the functioning of modality-specific perceptuomotor representations, without the involvement of representations of a more abstracted (conceptual) nature.

But can this account explain the size-brightness congruity effect observed in Experiment 1? In light of the correlation between substance brightness and portion size in the substances being sampled, might what appears to be an interaction between key size and brightness instead be an interaction between portion size and key size? This seems unlikely given that in Experiment 1 an interaction between portion size and key size itself was not observed.

Switching to edibility as the basis for classifying named substances in Experiment 2 provides a different, and arguably better, way of assessing the perceptuomotor simulation account of the size-brightness congruity effect. For two reasons the switch helps ensure that, while brightness remains relevant to the classification decision (albeit less directly than in Experiment 1) because of its diagnostic value regarding edibility, portion size also becomes relevant to the classification decision.

First, for the current sample of named substances, if not for substances more generally, the contrast between white and black is diagnostic of edibility (and the most obvious visual difference between flour and soot, and possibly the only visual difference, is their surface brightness). The diagnostic potential of surface brightness for the determination of edibility was strongly in evidence when the first author asked 71 undergraduate students to list, in 2 minutes, as many white (bright) and black (dark) things people can eat/drink, and not eat/drink, in any order. The outcome was clear, whereas bright substances were 3.8 times more likely to be edible than inedible, dark substances were 4.2 times more likely to be inedible than edible. On this basis it is reasonable to expect participants in the edibility classification task to make reference to brightness when arriving at a classification decision, assuming information about brightness is available in the perceptuomotor simulation.

Second, typical portion size also is relevant to the edibility classification decision for reasons relating to the sizes of hands and mouths. Only substances coming in (smaller) portion sizes appropriate for hands and mouths are likely to be edible, or to be thought of as having the potential to be edible. Again, therefore, it is reasonable to expect participants in the edibility classification task to make some reference to typical portion size when arriving at a classification decision, assuming information about portion size is available in the simulation.

But will information about brightness and portion size be incorporated in perceptuomotor simulations of the named substances? If task relevance is important for this, then for the reasons just given this would seem likely. But in fact there is evidence that perceptuomotor simulations automatically incorporate information about all features of named items (e.g., Connell & Lynott, 2009; Joseph & Proffitt, 1996; Wickens, Reutener, & Eggemeier, 1972). For example, when people hear the sentence *Susan liked it when her grandmother wore her hair up,* visual representations of both the typical colour of a grandmother's hair (grey), and the typical colour of hair in general (brown), are activated (Connell & Lynott, 2009). And, when people are presented with the sentence *Jane tasted the tomato before it was ready to eat*, not only is the atypical colour green activated, so also is the typical colour red, and just as strongly as it is activated by the sentence *Jane tasted the tomato when it was ready to eat.* As a final example, when people are presented with successive triplets of words to remember over a short period, release from proactive interference reveals their implicit sensitivity to the colours associated with the words. For example, when presented with the substance names *milk, chalk*, and *salt,* people automatically encode the fact that all the substances are white (Wickens, Reutener, & Eggemeier, 1972). In summary, it seems likely that information about both the brightness and typical portion size of a named substance will be represented in a perceptuomotor simulation, in which case interactions between brightness and key size, and between portion size and key size, would be expected. The latter interaction will be responsible for the former interaction because of the association between brightness and portion size, thereby denying the existence of a size-brightness congruity effect. With the outcome of Experiment 1 in mind, Experiment 2 was conceived to allow these issues to be explored.

Method

Participants

Twenty-seven Lancaster University students (19 females and eight males) aged between 18 and 29 (mean age = 20.5 years) volunteered to participate in the experiment in exchange for payment. All but one of the participants were right-handed by self-report. None of the participants had taken part in Experiment 1.

Task, Materials, and Apparatus

The task, materials, and apparatus were identical to those described in Experiment 1, except that participants were now required to decide whether the physical referent of each of the 10 stimulus words is edible or inedible (edible = *milk*, *sugar*, *salt*, *flour*, *yoghurt*; inedible = *coal, soil*, *ink*, *tarmac*, and *soot*). Participants also received a questionnaire containing three questions designed to assess the likelihood that they were referring to the brightness of the substances when making their classification decision, either using this (rather than edibility) as the basis for classifying the words, or using it as a secondary source of information to confirm the correctness of their provisional edibility decision.

Design and Procedure

The experimental design was identical to that described in Experiment 1, except that Edibility (edible vs. inedible) replaced Brightness (bright vs. dark) as a within-participant factor.

Procedure

The procedure was identical to that described in Experiment 1. However, rather than have participants classify the words according to the brightness of their referents (bright vs. dark), they were now asked to classify them according to the edibility of their referents (e.g., *flour* = edible, *coal* = inedible). Half of the participants (12 right-handers and one left-hander) were required to press the left-hand key whenever an item was edible, and the right-

hand key whenever it was inedible, and half (14 right-handers) were assigned to the opposite

edibility-hand mapping (i.e., right-hand key for edible and left-hand key for inedible).

The questionnaire was administered immediately after participants had completed the

speeded edibility classification task. It contained three questions designed to examine any

explicit classification strategies participants had adopted to help them perform the speeded

edibility classification task. The main goal of the questionnaire was to ensure that

participants had not reframed the task instructions to replace *edibility* with *brightness,* and

then used this distinction as the basis on which to classify the words. To this end,

participants were asked: how they decided whether an item was edible or inedible; whether

they had noticed anything about the edible and inedible words, and what this was; and,

whether or not they had used what they noticed to help them classify the words. Participants

were asked to answer all the questions as fully and honestly as possible in the order in which

they appeared in the booklet. They were instructed not to look ahead to later questions before

answering the previous ones, and not to change an answer once they had seen the upcoming

questions. Participants recorded all of their answers by hand. There was no time limit for

completing this task.

Results

Participants' explicit classification strategies were revealed by their responses to the

three questions in the questionnaire. Of particular interest was whether or not they indicated

noticing that all the edible items were bright and all the inedible items were dark, and

whether or not they had used this distinction to help them categorise the words during the

speeded edibility classification task. Five participants whose answers to at least one of these

questions indicated they had referred to the brightness of the substances when classifying

them were excluded from the analyses at this point. Only the data from the remaining 22

participants, who made no reference to the difference in the brightness of the edible and inedible words, are included in the analyses reported below.

The data were the accuracy and speed of responses to the words. Accuracy levels and mean observed correct RTs (after replacing excessively long RTs with cut-off values set at 2.5 *SD*s above a participant's mean RT) obtained for the bright and dark words, calculated separately for the small and big key, are shown in Table 2.

Table 2 Mean correct RT (*SE* in parentheses) and *p*(correct) for each of the edible (bright) and inedible (dark) words in Experiment 2 according to whether its edibility required a response on a 'congruently' sized key or an 'incongruently' sized key, where 'congruency' was defined in relation to the brightness of the items, rather than their edibility.

Test word	'Congruent'	'Incongruent'	RT difference ('inc'-'con')
bright			
milk	692 (23) 0.98	**633 (15) 0.99**	-59
sugar	695 (21) 0.98	**633 (13) 1.00**	-62
salt	765 (27) 0.92	**736 (19) 0.94**	-29
flour	737 (23) 0.97	**666 (17) 0.99**	-71
yoghurt	734(31) 0.99	**625 (16) 1.00**	-109
dark			
coal	**676 (19) 1.00**	725 (25) 1.00	49
soil	**670 (17) 0.99**	726 (23) 0.97	56
ink	**720 (15) 0.98**	755 (24) 0.96	35
tarmac	**645 (14) 1.00**	714 (22) 0.99	69
soot	**673 (22) 0.98**	706 (21) 0.99	33

Note. Normal and bold text relate to trials on which the small key and

the big key needed to be pressed, respectively

Response Accuracy

The overall mean level of response accuracy was 98.1% (*SD* = 4.4%). The mean

percentage of correct responses was not significantly higher for 'congruent' trials (98.0%)

than for 'incongruent' trials (98.3%), Wilcoxon Signed Ranks Test *z* = -0.95, *p* < .17, one-

tailed, where 'congruency' was defined in relation to the brightness of the items, rather than

their edibility.

Response Speed

The analysis replicated that adopted in Experiment 1, though now the decision being

made concerned the edibility of the named items, rather than their brightness. Nevertheless,

the factor distinguishing the two sets of words continues to be referred to as brightness.

There was now a significant WORD LENGTH X KEY SIZE interaction, $\chi^2(1) =$

16.14, *p* = .00006, with the congruence between these two factors raising response speed by

0.03 decisions/sec (*SE* = 0.009) for every additional letter in a word (see Figure 6). There

was not a significant PORTION SIZE X KEY SIZE interaction, $\chi^2(1) = 0.004$, *p* = .95. Nor

was there a significant BRIGHTNESS X KEY SIZE interaction, $\chi^2(1) = 0.41$, *p* = .52.

Indeed, inspection of the coefficients for the effect of the congruence between key size and

brightness, by word, indicated the reverse trend, with congruence appearing to have a

negative impact for seven of the ten words (coefficient range = -0.022 to 0.018 decisions/sec,

mean coefficient = -0.006, *SD* = .012). The observed effect of the congruence between

brightness and key size was -8 ms, reflecting a minor slowing of the observed mean RT from

699 to 707 ms.
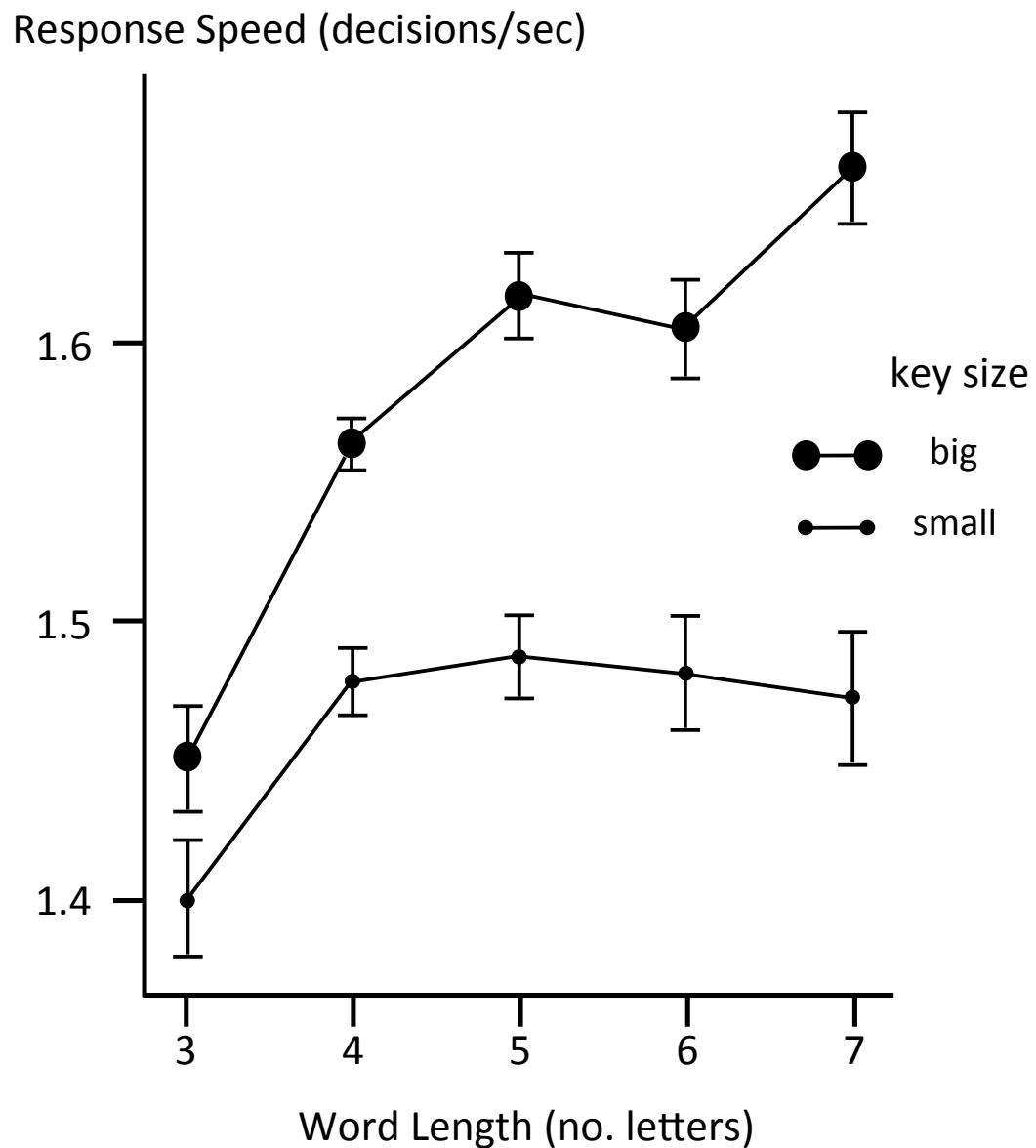
Figure 6 Experiment 2: Edibility classification. Mean observed response speed (decisions/sec) according to the length of the word being classified and the size of the key on which the edibility classification decision is being communicated (bars indicate *SE*s). Note the trend for response speed on the big key to increase with word length, relative to responses on the small key.

Discussion

The results of Experiment 2 confirm the predictions based on P. Walker et al.'s understanding of cross-sensory correspondences and the congruity effects they induce (see for example, L. Walker, P. Walker, & Francis, 2012; P. Walker & L. Walker, 2012a; P. Walker, 2012a, b). Specifically, the results support the proposal that the correspondence-induced size-brightness congruity effect is mediated by a subset of representations of size and brightness as conceptual dimensions accessible to stimuli across modalities, whether the particular feature values on these dimensions are specified verbally or perceptually. When the basis for classifying the named substances was switched from their brightness to their edibility the size-brightness congruity effect observed in Experiment 1 was no longer in evidence. This was predicted on the grounds that the switch to edibility would mean that the classification decisions were based on representations separate from those dealing with the feature dimensions involved in cross-sensory correspondences, thereby isolating performance from the factors responsible for the congruity effect.

Other aspects of the results from Experiment 2 offer additional support for the proposed explanation of the size-brightness congruity effect observed in Experiment 1, primarily by undermining alternative explanations. Because everything bar the conceptual basis for classifying the words was kept in place for Experiment 2, the absence of a congruity effect confirms that any features of the words themselves that might have remained inadequately controlled in Experiment 1 were not responsible for the size-brightness congruity effect. This rules out alternative explanations for the effect based on direct cross-domain mappings between the perceptual features of the words (e.g., their visual size, or their auditory pitch) and the perceptual features of the keys (most notably their haptic size). Also placed in doubt are explanations of the size-brightness congruity effect that draw exclusively on modality-specific sensory-perceptual representations, including perceptuomotor simulations. To the extent that the same perceptuomotor simulations would have been

activated by the visual and haptic stimuli in both experiments (i.e., despite the switch in classification decision), the same congruity effects should have been observed. However, a size-brightness congruity effect was observed only in Experiment 1. Furthermore, a congruity effect based on the interaction between typical substance portion size and key size that is predicted by the perceptuomotor simulation account, and that provides an alternative account of the size-brightness congruity effect, was not observed in either experiment. In Experiment 2 an interaction between typical portion size and key size was not observed despite the greater task relevance of portion size to the edibility classification decision needing to be made, and despite typical portion size correlating with brightness in the sample of substances under investigation.

Instead, and confirming the sensitivity of both experiments to interactions involving key size, Experiment 2 revealed a congruity interaction between the size (length) of the word being classified and key size, despite the fact that neither could inform the classification decision being made. It is important to note that this interaction could not explain the size-brightness congruity effect because the names of bright substances were longer, rather than shorter, than the names of dark substances (i.e., according to which bright substances should have been classified more easily on the bigger of the two keys). Though the theoretical significance of this enhanced sensitivity to features of the words themselves, rather than to characteristics of their referents, remains to be determined, it is clear that switching the classification decision away from a feature dimension involved in cross-sensory correspondences changed the type of representation informing the classification decision and, as a result, the factors impacting on performance. This is further indication that, as a consequence of the switch from brightness to edibility, the processes culminating in a classification decision ceased to make contact with the feature dimensions involved in cross-sensory correspondences.

There are other difficulties for any accounts of the size-brightness congruity effect based exclusively on interactions among sensory-perceptual representations. First, according to several researchers (e.g., Chiou & Rich, 2012; Gallace & Spence, 2006; Marks, 1974, 1987, 1989; Marks, Szczesiul, & Ohlott, 1986; Melara and Marks, 1990), such interactions should be tied to the absolute (context-insensitive) values of perceptual features, rather than their relative (context sensitive) values. However, it is the latter that are functionally important in the size-brightness congruity effect (L. Walker & P. Walker, 2015, and see above). Second, and linking with the functional significance of relative coding, P. Walker and L. Walker (2012) provide evidence for the involvement of post-categorical representations of size and brightness in a situation in which contrasting levels of brightness are specified non-verbally. They show, for example, how the size-brightness congruity effect they observed originated at levels of processing subsequent to the context-dependent (task dependent) brightness classification of each stimulus (see above). It is not clear how a perceptuomotor simulation account would explain the functional significance of the relative context-sensitive coding of stimulus features.

Finally, there is the general point, made by Melara and Marks (1990) and echoed elsewhere (e.g., L. Walker, P. Walker & Francis, 2012; P. Walker & L. Walker, 2012), that if the interactions responsible for cross-sensory correspondences involve modality-specific representations, then there should be no need for the same correspondences to emerge regardless of the modality through which they are probed, and no need for transitivity to be observed when switching between modalities. For example, while both high auditory pitch and visual brightness might be associated with visual smallness, high auditory pitch need not associate with visual brightness. And yet we see clear evidence for such transitivity when the same core set of correspondences emerge whichever sensory channels are used to probe them (L. Walker, P. Walker & Francis, 2012; P. Walker & L. Walker, 2012).

References

Baayen, R.J., Davidson, D.J. and Bates, D.M. (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language 59,* 390–412.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B, 364,* 1281-1289.

Brown, R & Nuttall, R. (1959). Method in phonetic symbolism experiments. *Journal of Abnormal and Social Psychology, 59,* 441-445.

Chiou, R. & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception, 41,* 339-353.

Chiou, R., Stelter, M., & Rich, A. N. (2013). Beyond colour perception: Auditory-visual synaesthesia induces experiences of geometric objects in specific locations. *Cortex, 49,* 1750-1763.

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.

Connell, L. & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin and Review, 16,* 573-577.

Connell, L. & Lynott, D. (2013). Flexible and fast: Linguistic shortcut affects both shallow and deep conceptual processing. *Psychonomic Bulletin and Review, 20,* 542-550.

Connell, L. & Lynott, D. (2014). Principles of representation: Why you can't represent the same concept twice. *Topics in Cognitive Science,* 1-17.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science, 24,* 613-621.

Eitan, Z. & 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.

Gallace, A. & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception and Psychophysics, 68,* 1191-1203.

Goldstein, H. (1994). Multilevel cross-classified models. *Sociological Methods and Research, 22,* 364–375.

Hauk, O. & Pulvermuller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping, 21,* 191-201.

Hornbostel, E. M. von. (1931). Über Geruschshelligkeit. *Pfugers Archiv fur die Gesamte Psysiologie, 227,* 517-538.

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

Joseph, J. E. & Proffitt, D. R. (1996). Semantic versus perceptual influence of color in object recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22,* 407-429.

Karwoski, T. F. & Odbert, H. S. (1938). Color-music. *Psychological Monographs, 50(2),* 1-60.

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

Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior, 10,* 427-429.

Kussner, M. B. & Leech-Wilkinson, D. (2013). Investigating the influence of musical training on cross-modal correspondences and sensorimotor skills in a real-time drawing task. *Psychology of Music.* DOI: 10.1177/0305735613482022

Louwerse, M. & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science, 35,* 381-398.

Lunghi, C. & Alais, D. (2013). Touch interacts with vision during binocular rivalry with a tight orientation tuning. *PLoS ONE, 8(3),* e58754. doi: 10.1371/journal.pone.0058754

Lunghi, C., Binda, P., & Morrone, M. C. (2010). Touch disambiguates rivalrous perception at early stages of visual analysis. *Current Biology, CB, 20(4),* R 143-R144. doi: 10.1016/j.cub.2009.12.015

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: 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. (1989). On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 586-602. doi: 10.1037/0096-1523.15.3.586

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: MIT Press.

Marks, L. E., Szczesiul, R., & Ohlott, P. (1986). On the cross-modal perception of intensity. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 517-534. doi: 10.1037/0096-1523.12.4.517

Martino, G. & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception, 28,* 903-923.

Melara, R. D. & Marks, L. E. (1990) Processes underlying dimensional interactions: Correspondences between linguistic and non-linguistic dimensions. *Memory and Cognition, 18,* 477-495.

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

Orchard-Mills, E., Alais, D., & van der Burg, E. (2013). Cross-modal associations between vision, touch, and audition influence visual search through top-down attention, not bottom-up capture. *Attention, Perception and Psychophysics, 75,* 1892-1905. doi: 10.3758/s13414-013-0535-9

Orchard-Mills, E., van der Burg, E., & Alais, D. (2013). Amplitude-modulated auditory stimuli influence selection of visual spatial frequencies. *Journal of Vision, 13(3),* 6, 1-17. doi: 10.1167/13.3.6

Osgood, C.E., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning.* Urbana, IL: University of Illinois Press

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.

Pulvermuller, F., Harle, M., & Hummel, F. (2001). Walking or talking?: Behavioral and

    neurophysiological correlates of action verb processing. *Brain and Language, 78,*

    143-168.

Pulvermuller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in

    action word recognition. *Journal of Cognitive Neuroscience, 17,* 884-892.

Rhodes, R. (1994). Aural images. In L. Hinton, J. Nichols, and J. J. Ohala (Eds.), *Sound*

    Symbolism (pp. 276-292). Cambridge, UK: Cambridge University Press.

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

    239–255.

Smith, L. B. & Sera, M. D. (1992). A developmental analysis of the polar structure of

    dimensions. *Cognitive Psychology, 24,* 99-142.

Solomon, K. O. & Barsalou, L. W. (2004). Perceptual simulation in property verification.

    Memory and Cognition, 32, 244-259.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, &*

    Psychophysics, 73, 971-995.

Thompson, P. D. & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded

    function. *Quarterly Journal of Experimental Psychology, 64,* 2392-2404.

University of Oxford. (2010). British National Corpus. Retrieved February 3, 2012 from

    http://www.natcorp.ox.ac.uk/index.xml

Walker, P. (2012a). Cross-sensory correspondences and crosstalk between dimensions of

    connotative meaning: Visual angularity is hard, high-pitched, and bright. *Attention,*

    Perception, & Psychophysics, 74, 1792-1809.

Walker, P. (2012b). Cross-sensory correspondences and naïve conceptions of natural

    phenomena. *Perception, 41,* 620-622.

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, P. & Walker, L. (2012). Size-brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning. *Attention, Perception, & Psychophysics, 74(6),* 1226-1240.

Walker, L. & Walker, P. (2015). Cross-sensory mapping of feature values in the size-brightness correspondence can be more relative than absolute. *Submitted*

Walker, L., Walker, P., & Francis, B. J. (2012). A common scheme for cross-sensory correspondences across stimulus domains. *Perception, 41(10),* 1186-1192.

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

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.

Wickens, D. D., Reutener, D. B., & Eggemeier, F. T. (1972). Sense impression as an encoding dimension of words. *Journal of Experimental Psychology, 96,* 301-306.

Zigler, M. J. (1930). Tone shapes: A novel type of synaesthesia. *Journal of General Psychology, 3,* 277-286.

Appendix

The word-level features possessed by words in the bright and dark sets were assessed in various ways. With regard to the nature of the vowels and consonants in the words, reference was made to the current version of the International Phonetic Alphabet (IPA). The average position of the vowels was scored with reference to which part of the tongue is raised during articulation (*front* = 1; *central* = 2; *back* = 3). This variation is regarded as translating into systematic variation in the acoustic frequency of the second formant of the vowel sound, running from *front* = high frequency to *back* = low frequency. The average height of the vowels was scored with reference to how high the tongue is raised during articulation (*close/high tongue* = 1; *close-mid* = 2; *open-mid* = 3; *open/low tongue* = 4). This variation is regarded as translating into systematic variation in the acoustic frequency of the first formant of the vowel sound, running from *close/high tongue* = low frequency to *open/low tongue* = high frequency. The visual angularity of each word's letter forms was assessed by asking a group of 25 undergraduate students to rate each word on a 6-point Likert scale, running from *very angular* = 1 to *very rounded* = 6. The words appeared (i.e., font/size/colour) exactly as they had in the classification task. The IPA was consulted to identify which consonants in the words were plosive in nature. Finally, the portion sizes in which the named substances are typically encountered was assessed by asking a different group of 28 undergraduate students to rate each word on a 6-point Likert scale, running from *very small* = 1 to *very big* = 6. The mean values for each word-level feature for each set of words are presented in Table 3.

Table 3   Mean values (with *SD*s in parentheses) for word-level features having the potential to interact directly with response key size and, therefore, with the potential to provide an alternative explanation for the size-brightness congruity effect

Word-level feature	Bright	Dark
average position of vowels	2.22 (0.58)	2.10 (0.38)
average height of vowels	2.42 (0.72)	2.20 (0.92)
size of words (number of letters)	5.00 (1.22)	4.20 (1.10)
visual angularity of word forms	3.39 (1.30)	3.49 (1.57)
percentage consonants that were plosives	33.3 (23.6)	43.3 (25.3)
portion size of named substances	2.11 (0.64)	2.94 (1.04)