

RUNNING HEAD: CONCRETE AND ABSTRACT CONCEPTS

The Representation of Concrete and Abstract Concepts:

Categorical vs. Associative Relationships

Jingyi Geng and Tatiana T. Schnur

Department of Psychology, Rice University

In press, *Journal of Experimental Psychology: Learning, Memory, & Cognition*

Please address correspondence to:

Tatiana T. Schnur, Ph.D.

Department of Psychology

Rice University, USA

Houston, Texas 77005

E-mail: ttschnur@rice.edu

Phone: + 1 713 348 5054

This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Abstract

In four word-translation experiments, we examined the different representational frameworks theory (Crutch & Warrington, 2005, 2010) that concrete words are represented primarily by category whereas abstract words are represented by association. In our experiments, Chinese- English bilingual speakers were presented with an auditory Chinese word and three or four written English words simultaneously and asked to select the English word that corresponded to the auditory word. For both abstract and concrete words, higher error rates and longer response times were observed when the English words were categorically or associatively related compared to the unrelated conditions and the magnitude of the categorical effect was bigger than the associative effect. These results challenge the different representational frameworks theory and suggest that although category and association are important for representing abstract and concrete concepts, category plays a greater role for both types of words.

Keywords: Concrete and abstract words; Semantic processing; Semantic categorical and associative effects; Bilingual translation

Investigating the representations of concrete (e.g., *mouse*) and abstract words (e.g., *future*) has important implications for understanding how our memory of factual information and general knowledge of the world (i.e., semantic memory) is represented. A number of studies show that concrete words are more easily recognized and memorized in various cognitive tasks (e.g., lexical decision, word reading) compared to abstract words (e.g., James, 1975; Kroll & Merves, 1986; Paivio, 1971; Strain, Patterson, & Seidenberg, 1995; see Paivio, 1991, for a review). More importantly, individuals with semantic memory deficits show a selective impairment of concrete knowledge with relatively preserved abstract knowledge or vice versa (e.g., Breedin et al., 1994; Sirigu et al., 1992; Warrington, 1975; Warrington & Shallice, 1984), suggesting that concrete and abstract concepts are represented in a different way. Concepts may be represented by taxonomic categories defined by the overlap in the semantic features between concepts (e.g., *ducks* and *pigeons* are in the *bird* category, as they have wings, feathers and can fly), or alternatively, by the links between concepts that tend to co-occur in language or which form familiar scenes or events (by association, e.g., “The *mouse* ate the *cheese*”). Most research has focused on concrete concepts, showing that semantic category and association both play a key role in representing concrete concepts (e.g., Kalénine et al., 2009; Sachs, Weis, Krings, Huber, & Kircher, 2008; Sachs, et al., 2011; Schwartz, et al., 2011; for reviews, see Hutchison, 2003; Mahon & Caramazza, 2009; Martin, 2007), but the representations of abstract concepts are less explored. The recent different representational frameworks theory proposes that category and association are important for representing both concrete and abstract words but concrete words rely more on category whereas abstract words rely more on association (Crutch & Warrington, 2005, 2010).

Although this theory has support from neuropsychological studies of individuals with aphasia (e.g., Crutch & Warrington, 2005, 2007, 2010) and a few studies with healthy people (Crutch, Connell, & Warrington, 2009; Crutch & Jackson, 2011), there exist conflicting results (Hamilton & Coslett, 2008; Hamilton & Martin, 2011; Zhang, Han, & Bi, 2012) in addition to several methodological issues (e.g., different category definitions for concrete and abstract words; unmatched semantic distance; non-comparable tasks between patients and healthy participants) which are addressed below. These methodological issues render the results difficult to interpret. Thus, the aim of this study was to test predictions not previously examined in healthy people from the different representational frameworks theory of the representation of concrete and abstract concepts (Crutch & Warrington, 2005, 2010).

The first evidence that abstract and concrete conceptual knowledge is represented differently within categorical and associative relationships comes from behavior in individuals with brain damage resulting in aphasia. Crutch and Warrington (2005) tested an aphasic patient (A.Z.) with a spoken-word to written-word matching task, where the patient identified the word she heard among four written words, which were all either concrete or abstract words. For concrete words, A.Z. made more errors when the four words were categorically related (e.g., *goose, crow, sparrow, pigeon*) vs. unrelated (e.g., *goose, melon, pullover, biscuit*), but showed no difference when the four words were associatively related (e.g., *farm, cow, tractor, barn*) vs. unrelated (e.g., *sailor, shelf, farm, oven*). An opposite pattern was observed for abstract words (i.e., more errors when the words were associatively related (e.g., *comedy, joke, laugh, funny*) vs. unrelated and no difference between the categorically related (synonyms or near-synonyms, e.g., *scant, bare, sparse, deficient*) and

unrelated conditions). Three additional patients showed a similar pattern in later studies (I.R.Q.: Crutch et al., 2006; R.O.M.: Crutch & Warrington, 2007; F.B.I.: Crutch & Warrington, 2010). However, subsequently the Crutch and Warrington (2005, 2007, 2010) results were not fully replicated where three aphasic patients (UM-103, H.A. and D.Z.; Hamilton & Coslett, 2008; Hamilton & Martin, 2010) made more errors in both categorically and associatively related conditions compared to the unrelated condition (in the same task as Crutch and Warrington, 2005), suggesting that category and association are equally important for representing concrete and abstract words. Based on these results, Crutch and Warrington (2010, pp. 49 – 50) suggested that the inconsistent patient evidence depends on whether impairments to category and association systems are equivalent. If aphasic speakers have equivalent impairments for category and association systems (e.g., the patients reported in Crutch & Warrington, 2005; Crutch et al., 2006; Crutch & Warrington, 2007; and Crutch & Warrington, 2010) they should show greater errors selecting a target from amongst categorically related concrete words (but not when they are associatively related) and greater errors selecting a target from amongst associatively (but not categorically) related abstract words. In contrast, similar categorical and associative interference effects are observed for concrete and abstract words if aphasic speakers have unequal or selective damage to category and association systems (UM-103, Hamilton & Coslett, 2008; H.A. and D.Z., Hamilton & Martin, 2011). Because Crutch and Warrington (2010) argue that the performance of aphasic patients depends on the degree of damage to the category and association systems, testable predictions cannot be generated to distinguish between accounts of semantic representation where category and association are equally important or differentially important for the

representation of concrete and abstract words.

However, behavior in healthy speakers whose semantic systems are intact provides converging evidence to disambiguate whether concrete and abstract concepts depend equally on categorical and associative relationships. Unfortunately, these data to date are inconsistent. In a semantic odd-one-out task, healthy participants identified a word unrelated to other words more quickly when concrete words were in categorical vs. associative arrays and more quickly when abstract words were in associative vs. categorical arrays (Crutch, Connell, & Warrington, 2009; Crutch & Jackson, 2011). In contrast, in an oral translation task, although Chinese-English bilingual speakers translated concrete Chinese words to English more slowly in categorically related than unrelated blocks (with no difference between the associatively related and unrelated blocks), participants translated abstract words more slowly when the words were blocked by either association or category compared to unrelated blocks (Zhang, Han & Bi, 2012). These latter results suggest that in contrast to the different representational frameworks theory, concrete words are represented only by category whereas abstract words are equally represented by both association and category. Because the results from healthy people and brain-damaged patients are so mixed (potentially due to methodological confounds, described below) it is unclear the degree to which category and association are involved in the representation of concrete and abstract concepts.

Methodological confounds

There are several methodological issues in the previous studies which cloud clear interpretation of current results, which we address in our experiments. First, different definitions are applied for what constitutes a taxonomic category for concrete vs. abstract

words. According to Crutch and Warrington's (2005, 2007, 2010) classification, categorically related concrete words are defined within a taxonomic structure (e.g., *geese* and *pigeons* are in the "*bird*" category) but categorically related abstract words are synonyms or near-synonyms (e.g., *look*, *see*). Because taxonomic category was defined differently for concrete and abstract words in previous studies (Crutch & Warrington, 2005, 2007, 2010; Crutch et al., 2009; Crutch & Jackson, 2011; Hamilton & Coslett, 2008; Hamilton & Martin, 2011; Zhang et al., 2012), it is unclear whether the contrast in the categorical effects between concrete and abstract words in those studies truly reflects the different representation of concrete and abstract words. In contrast to the claim that abstract words cannot be defined within a taxonomic structure but concrete words can (Crutch et al., 2009, p. 1387), Verheyen, Stukken and De Deyne (2011) argued that abstract and concrete words have similar category structure in terms of the overlap of features, including features such as typicality, exemplar generation, etc. To address the confound of differently defined taxonomic categories for abstract vs. concrete words, in our study we employed the Verheyen et al. category structure to create the same category definitions for concrete and abstract words.

Second, in the patient studies, the stimuli in the categorically and associatively related conditions were not well matched for semantic distance. If semantic distance is not controlled for, interference effects might reflect the difficulty in searching among words with close semantic distance (e.g., Campanella & Shallice, 2011), instead of the differential dependency on category vs. association in concrete and abstract words. Although Crutch and Warrington (2007) collected word-set ratings for category and association strength (on a scale from associatively related (-3), unrelated (0), to categorically related/ synonym (+3)), this rating

scale did not accurately measure semantic distance between concepts because participants could only choose between one of the two relationships, but could not indicate the case where the word sets were both categorically and associatively related (for a similar argument, see Zhang et al., 2012). Employing a more objective semantic distance measurement, Latent Semantic Analysis (LSA, Landauer, Foltz, & Laham, 1998), we found that for abstract words, the semantic distance in the Crutch and Warrington (2007)'s associatively related condition (.33) was closer than the categorically related condition (.24) ($t(34) = 1.96, p = .06$), which may explain the higher error rate observed in the associatively vs. categorically related conditions for these materials (Crutch & Warrington, 2007, 2010) negating the evidence of a greater role of association in representing abstract words. Therefore, to address the semantic distance confound we used LSA measures to balance our materials for semantic distance. In addition, to verify the validity of the categorical and associative relationship manipulation, we also collected separate ratings for categorical and associative strength and used identical instructions for both concrete and abstract words.

Third, among the three studies with healthy people, two did not test whether categorical and associative effects emerge for both concrete and abstract words (Crutch, et al., 2009; Crutch & Jackson, 2011), as the nature of the semantic odd-one-out task (i.e., identify an unrelated word from a categorically or associatively related word array), does not allow an unrelated condition as baseline. These two studies (Crutch, et al., 2009; Crutch & Jackson, 2011) directly compared the response times in the categorically related word arrays with those in the associatively related word arrays and then concluded that abstract words are represented only by association and concrete words only by category. Without a baseline

condition, it is unclear whether both categorical and associative effects (i.e., the difference between related and unrelated conditions) are observed for both concrete and abstract words, which is a critical prediction if these are the principles delineating concrete and abstract concepts. In our study, we address both predictions from the different representational frameworks theory where first, categorical and associative effects should be observed for both concrete and abstract words, and second, we should observe a larger categorical effect for concrete words but a larger associative effect for abstract words.

Lastly, although Zhang et al. (2012) included an unrelated condition as baseline in an oral translation task, interference effects in such a production task are difficult to interpret because they may reflect interference at a lexical level (e.g., Bloem & La Heij, 2003, 2004). Specifically, because oral translation involves not only conceptual activation but also lexical selection for production, the categorical interference effects observed for both concrete and abstract words in the translation task may reflect lexical selection by competition instead of, or in addition to conceptual identification (e.g., Abdel Rahman & Melinger, 2009; Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992; Levelt, Roelofs, & Meyer, 1999; but see Mahon, Costa, Peterson, Vargas, & Caramazza, 2007). As a result, these results too are difficult to interpret with regards to the representation of concrete and abstract words. To address this concern, in our study we adopted a comprehension task to avoid the confound of lexically-based interference in production.

Current Study

The purpose of this study was to test two predictions from the different representational frameworks theory of concrete and abstract words (Crutch & Warrington, 2010) to provide

converging evidence in healthy participants while addressing the methodological issues summarized above. In addition, to better compare the performance of healthy participants with performance of aphasic patients (e.g., Crutch & Warrington, 2005, 2010) we adopted a speeded word translation task which allows analysis of not only the response times but also the error rates in healthy participants (see Campanella & Shallice, 2011 for a similar paradigm which produced response time and error patterns in healthy participants similar to semantic aphasia patients). Because both naming/comprehending pictures and translating words is slower when the pictures/words are grouped together in the same semantic category vs. grouped in different categories (e.g., Kroll & Stewart, 1994; Campanella & Shallice, 2011), this suggests that translation tasks capture access to conceptual representations in a similar way as seen in other language comprehension tasks (e.g., picture matching). Hence, word translation is an appropriate paradigm to study access to semantic representations.

In our word translation task, Chinese- English bilingual participants selected an English word that corresponded to the Chinese word they heard, from an array of four English words categorically related, associatively related or unrelated. We predicted that if abstract and concrete words depend (differentially so) on categorical and associative relationships (i.e., Crutch & Warrington, 2010) then the categorical and associative effects should be observed for both concrete and abstract words, but for concrete words, the size of the categorical effect should be bigger than the associative effect whereas for abstract words, the opposite should be observed. In contrast, if abstract and concrete knowledge depend equally on categorical and associated knowledge (i.e. Hamilton & Coslett, 2008; Hamilton & Martin, 2010) we predicted equal associative and categorical effects for concrete and abstract concepts. We

tested these predictions using concrete words (Experiment 1) and abstract words represented in categories either using synonyms (to replicate Crutch and colleagues, Experiment 2) or abstract taxonomically represented categories from Verheyen et al. (2011) (Experiment 3). Lastly, Experiment 4 directly compared concrete and abstract word comprehension in the same participants. By improving upon previous study designs and employing a new comprehension task, this study helps us to further understand the role of semantic relations in the representation of concrete and abstract words.

Experiment 1 – Concrete Words

Method

Participants. 32 undergraduate students (Age: 19 years \pm 1.2) at Rice University participated in this experiment. In this and the following experiments, all participants were Chinese-English bilinguals, with more than 10 years of fluent English and Chinese experience (English: 15 years \pm 3.3; Chinese: 18 years \pm 2.3). Participants received class credit for their participation. In accordance with the protocol approved by Rice University's Institutional Review Board, participants across all experiments gave their informed consent.

Materials. We used 64 concrete English words and their Chinese translations. Stimuli in the associatively related condition were taken from Crutch and Warrington (2005, Experiments 4 and 5; 2007, Experiment 2), where thirty-two words were used in the associatively related condition and regrouped to form the unrelated condition. In the categorically related condition, there were eight categories: clothing, animals, vegetables, food, fruits, birds, instruments, and vehicles. These categories have been previously shown to demonstrate semantic effects in a number of comprehension and production tasks (e.g.,

Caramazza & Costa, 2000, 2001; Estes & Jones, 2009; Mahon, Costa, Peterson, Vargas, Caramazza, 2007; McRae & Boisvert, 1998; Moss, Ostrin, Tyler, Marslen-Wilson, 1995) and are related by visual (i.e., birds have fur and wings) and function features (i.e., instruments can play music) among others. There were four items in each category. Thirty-two words were used in the categorically related and unrelated conditions (see Appendix A). Following Nelson and Schreiber (1992), we defined concrete words as words with concreteness ratings between 470-700 (ratings from the MRC psycholinguistic database, Coltheart, 1981). The average rating of concreteness for all concrete words was 607 (range 535 – 646). To verify the validity of the categorical and associative relationship manipulation, we collected categorical and associative ratings for concrete words in Experiments 1 and 4 from 12 participants and abstract words in Experiment 2 and 3 from another 12 participants. These 24 participants did not participate in any of the experiments. Following the instructions used in recent studies (e.g., Mirman & Graziano, 2012; Schwartz et al., 2011; Zhang et al., 2012), in the categorical rating session, participants were asked to “decide to what extent these four things are members of the same category”. In the associative rating session, participants were asked to “decide to what extent these four things co-occur in a situation or scene.” The associative ratings for the associatively related word arrays (average: 6.86) were significantly higher than the associatively unrelated word arrays (average: 1.25) ($t(14) = 63.31, p < .001$). Similarly, the categorical ratings for the categorically related arrays (average: 6.24) were significantly higher than the categorically unrelated arrays (average: 1.08) ($t(14) = 29.60, p < .001$). Importantly, the associative ratings for the associatively related arrays (average: 6.86) were significantly higher than the categorically related arrays (average: 3.88) ($t(14) = 7.19, p$

< .001). The categorical ratings for the categorically related arrays (average: 6.24) were significantly higher than the associatively related arrays (average: 2.57) ($t(14) = 19.90, p < .001$) (see Table 2).

The stimuli in the categorical conditions had higher imageability and concreteness ratings (Coltheart, 1981) than those in the associative conditions (imageability: $t(62) = 2.69, p = .01$; concreteness: $t(62) = 1.92, p = .06$), but did not differ in word frequency and word length (Coltheart, 1981) (see Table 1). Due to the fact that exactly the same words were used in the related (categorically or associatively) and unrelated conditions, any particular item effects (e.g., word frequency, word length, imageability) in the unrelated conditions are predicted to also occur in the related conditions regardless of semantic context (category or association). Thus, subtracting the unrelated from the related condition to reveal categorical/associative effects subtracts out any particular item differences, and leaves simple context effects (category or association). However, although we believe that item differences (categorical vs. associative) in our experiments did not influence the related vs. unrelated results, in order to dispel any concerns about between-item differences, in Experiment 4 we replaced several Experiment 1 items with new items (for concrete words) to match the variables known to impact word reading. We also included the variables influencing word reading as covariates in the data analyses (see footnote 2).

Given that previous studies showed an impact of semantic distance on response times in both comprehension and production paradigms (e.g., Campanella & Shallice, 2011; Vigliocco, Vinson, Damian, & Levelt, 2002), we used the LSA (Landauer, Foltz, & Laham, 1998; <http://lsa.colorado.edu>) as a measure of semantic distance. By averaging the LSA values for

all word pairs in each array, we calculated eight LSA averages for each condition (the same method applied for Experiments 2-4). The associatively and categorically related arrays were significantly closer in semantic distance than their corresponding unrelated arrays (associatively related vs. unrelated: $t(14) = 7.35, p < .001$; categorically related vs. unrelated: $t(14) = 4.47, p < .001$) but were not significantly different from each other (associatively vs. categorically related: $t < 1$). The semantic distance was not different between the two unrelated conditions either ($t < 1$).

All stimuli in this and following experiments were adapted to Chinese by direct translation. In order to verify that the Chinese-English translations in this and following experiments were transparent, we recruited an additional 18 English-Chinese bilingual participants who did not participate in Experiments 1-4 (Age: 20 ± 2.4 years; English/Chinese bilinguals for more than 10 years (English: $14 \text{ years} \pm 4.1$; Chinese: $19 \text{ years} \pm 3.1$)), to complete a translation evaluation task where we presented a written English word and three or four either categorically or associatively related written Chinese words from the three experiments. Participants chose a Chinese word from the related array to match the English word according to the meaning. We predicted that if the Chinese translations were confusing, they would make errors in choosing the Chinese word corresponding to the English word. If the error rate for a particular item was higher than 20%, we excluded that item from the data analysis across the four experiments. As a result, in Experiment 1, one item (*blouse*) in the categorically related condition and another item (*camp*) in the associatively related condition had more than 20% error, which we subsequently excluded from the data analysis.

---- Insert Table 1 about here ----

Design. The experiment consisted of four conditions, each with eight arrays. Each array consisted of four items. The conditions were as follows:

- a) *associatively related:* All four words were associated with each other (e.g., *farmer, cow, tractor, barn*)
- b) *associatively unrelated:* The four words were not associatively or categorically related to each other (e.g., *camp, oven, anchor, monkey*). The associatively unrelated arrays used the same items as the associatively related condition.
- c) *categorically related:* All four words were in the same semantic category (e.g., *skirt, jacket, pants, blouse*).
- d) *categorically unrelated:* All four words were not associatively or categorically related to each other (e.g., *pants, ship, eggplant, rabbit*). The categorically unrelated arrays used the same items as the categorically related condition.

Following the study design by Crutch and Warrington (2005, 2007, 2010), each condition formed a block in which the eight arrays repeated four times, totaling 128 trials per block and 512 in the entire experiment. The order of block presentation was counterbalanced across participants. We chose the blocked-cyclic experimental design across all four experiments to replicate the paradigm used in the previous five patient studies reviewed in the introduction (Crutch & Warrington, 2005, 2007, 2010; Hamilton & Coslett, 2008, Hamilton & Martin, 2011)¹.

Procedure. This experiment was operated on DMDX software (Forster & Forster, 2003). All participants were tested individually, seated at a distance of about 65 cm from the

computer screen. All responses were made on a standard PC keyboard. Testing was divided into four sessions: the language test, the familiarization session, the practice session, and the experimental session.

We used a language test to determine whether a participant was sufficiently fluent in both English and Chinese to participate in the experiment. During the language test, participants heard a Chinese word and were presented with an English word. Participants indicated whether or not the two stimuli were matched by selecting either a “yes” or “no” key. All the experimental stimuli were presented in this test. Participants had two seconds to respond. If accuracy were lower than 80%, the participant’s data were excluded.

Following other translation studies (e.g., Bloem & La Heij, 2003; Navarrete & Costa, 2009; Zhang et al., 2012), participants were given a familiarization session where the correct pairs of auditory and visual stimuli from the experiment were presented. Participants heard a Chinese word and saw an English word simultaneously. Each pair was only presented once. Following familiarization, a practice session consisted of four items not used in the experimental session, which were organized into one array. Each item was used as the target item four times, creating 16 trials during the practice session. Trials were run in the same way in both the practice and experimental sessions. Participants heard a Chinese word and were simultaneously presented with four English words. Each English word was randomly assigned a number from 1 to 4. Participants selected the English word that matched the Chinese word by pressing the corresponding number on the keyboard. Once the participant selected a response, or two seconds passed without a response, the next trial began (0 ms inter-stimulus interval). The same four English words were then pseudorandomly rearranged

and another Chinese word was presented from the same array. This procedure was repeated until all four Chinese words had been presented four times, each time utilizing the same array. The participant then moved on to the next array until the end of the block. At this point the participant was allowed a short resting period before continuing on to the next block. The duration of the experiment was 40 minutes.

Results and Discussion

All participants made less than 20% error on the language assessment test. Due to more than 20% errors for two items (i.e., *blouse*, *camp*) in the translation evaluation task, we excluded these two items from the data analysis. Response times were discarded from the analyses whenever: (a) an incorrect English word was chosen; (b) there was no response; or (c) response times deviated from a participant's mean by more than three standard deviations resulting in exclusion of 13% of the data. The words to be translated (i.e., target words) were considered as items in this and following experiments. The semantic context was a within-participant and between-item variable and relatedness was a within-participant and within-item variable. We used linear mixed-effects modeling (e.g., Baayen, Davidson, & Bates, 2008) for the response time analysis in this and following experiments implemented in JMP Pro 10. The response times were fit into linear regression models with random and fixed effects. Random factors included participants and items nested within semantic context. Fixed factors included semantic context (category vs. association), relatedness (related vs. unrelated), and the interaction between semantic context and relatedness. Logistic regression (e.g., Agresti, 2002) was employed for the error analyses in this and following experiments implemented in JMP Pro 10. Fixed factors were the same as the linear mixed-effects model.

An overview of the mean error rates, response times, and standard deviations for this experiment is presented in Tables 3 and 4. The differences in the response times and error rates between the associatively related vs. unrelated and categorically related vs. unrelated conditions for this experiment are presented in Figure 1.

---- Insert Table 3 about here ----

---- Insert Table 4 about here ----

---- Insert Figure 1 about here ----

In the error analysis², there was a significant main effect of semantic context ($\chi^2(1) = 91.56, p < .001$) with more errors in the associatively related/unrelated conditions (.12) than the categorically related/unrelated conditions (.09)³. The main effect of relatedness was also significant ($\chi^2(1) = 104.39, p < .001$) where participants made more errors in the categorically/associatively related conditions (.13) than the unrelated conditions (.08). There was a significant interaction between semantic context and relatedness ($\chi^2(1) = 17.98, p < .001$) where the error rate difference was larger in the categorically related vs. unrelated condition (.06), compared to the difference between the associatively related vs. unrelated condition (.03). The simple effect tests showed that error rates were significantly higher in the associatively related condition (.13) vs. the unrelated condition (.10) ($\chi^2(1) = 25.31, p < .001$) and the categorically related (.12) vs. the unrelated condition (.06) ($\chi^2(1) = 96.61, p < .001$).

In the response time analysis, there was a main effect of semantic context ($F(1, 60) = 16.94, p < .001$), where response times in the associatively related/unrelated conditions (1337 ms) were slower than those in the categorically related/unrelated conditions (1255 ms). The main effect of relatedness was significant ($F(1, 14130) = 132.56, p < .001$) where response

times in the related conditions (1322 ms) were slower than in the unrelated conditions (1270 ms). Similar to the response time analysis, there was a significant interaction between semantic context and relatedness ($F(1, 14130) = 23.77, p < .001$) where the size of the categorical effect (74 ms) was significantly larger compared to the associative effect (30 ms). The simple effect tests showed that for the associative effect, there was a significant difference between associatively related (1352 ms) and unrelated (1322 ms) conditions ($F(1, 14129) = 21.76, p < .001$), and response times were significantly longer for categorically related (1292 ms) vs. unrelated (1218 ms) items ($F(1, 14129) = 136.06, p < .001$).

In sum in Experiment 1, for both error and response time analyses, we observed significant associative and categorical interference effects, where the categorical interference effect was significantly larger than the associative effect for concrete concepts. This pattern suggests that for concrete concepts, categorically related concepts form closer neighbors compared to the associatively related concepts, resulting in greater difficulty to discriminate between categorically related concepts (for a similar argument, see p. 622, Crutch & Warrington, 2005). This result is consistent with the different representational frameworks theory, suggesting that concrete words are represented by both semantic category and association but rely more on category.

Experiment 2 – Abstract Words (synonyms)

In Experiment 2 we tested both associative and categorical effects for abstract words where the different representational frameworks theory (Crutch & Warrington, 2005, 2010) predicts significant associative and categorical effects, but the former larger than the latter. We defined category in abstract words by employing the category definition (i.e.,

near-synonyms) from previous studies (e.g., Crutch & Warrington, 2005, 2007, 2010; Hamilton & Martin, 2011). Due to the inconsistent results from previous studies with patients for abstract words (i.e., no synonym (categorical) effect in Crutch & Warrington, 2005, 2007, 2010; a significant synonym (categorical) effect in Hamilton & Coslett, 2008; Hamilton & Martin, 2011), it is critical to test whether the synonym effect is observed in healthy participants with the same materials. In Experiment 3 we tested the category structure for abstract words in a way more similar to that for concrete words (i.e., both taxonomically, e.g., *animals: cat and rabbit; sciences: physics and biology*; cf. Verheyen et al. 2011). Unless specified, Experiment 2 methods were identical to Experiment 1.

Method

Participants. 32 Chinese-English bilingual participants (Age: 19 years \pm 1.0) took part in this experiment, speaking both English and Chinese more than 10 years (English: 15 years \pm 3.2; Chinese: 18 years \pm 1.5). None participated in Experiment 1.

Materials. Experiment 2 consisted of 72 abstract English words and their Chinese translations. Abstract word stimuli and their division as associatively or categorically related were taken from a subset of materials used by Crutch and Warrington (2007, Experiment 2) and Zhang, et al. (2012, Experiment 2). Thirty-six words were used in both associatively related and unrelated conditions and the other 36 were used in both categorically related (synonyms) and unrelated conditions (see Appendix B). This experiment was originally run with all 72 words. However, based on Nelson and Schreiber (1992) to better control for degree of abstractness, prior to data analysis we eliminated words with concreteness ratings above 470 (ratings taken from the MRC psycholinguistic database, Coltheart, 1981).

According to this criterion, six abstract words (i.e., *bank, bite, disease, family, market, money*) were excluded from the data analysis. The average rating of concreteness for the remaining stimuli was 360 (range 241 – 470). Critically, the abstract words in Experiment 2 had significantly lower concreteness compared to the concrete items in Experiment 1 (concreteness: $t(124) = 19.67; p < .001$).

The associative and categorical ratings results (see Experiment 1 Methods) showed that the associative ratings for the associatively related word arrays (average: 6.07) were significantly higher than the associatively unrelated word arrays (average: 1.90) ($t(22) = 18.65, p < .001$). Similarly, the categorical ratings for the categorically related arrays (average: 5.33) were significantly higher than the categorically unrelated arrays (average: 1.45) ($t(22) = 14.26, p < .001$). Importantly, the associative ratings for the associatively related arrays (average: 6.07) were significantly higher than the categorically related arrays (average: 4.69) ($t(22) = 4.38, p < .001$). The categorical ratings for the categorically related arrays (average: 5.33) were significantly higher than the associatively related arrays (average: 3.58) ($t(22) = 5.80, p < .001$) (see Table 2).

The stimuli in the categorical conditions had lower imageability (Coltheart, 1981) than those in the associative conditions ($t(64) = 2.11, p = .04$). Words in the associative conditions had more letters than in the categorical conditions ($t(64) = 2.62, p = .01$) but they did not differ in word frequency (Coltheart, 1981) (see Table 1). To control for an interaction between these lexical factors and the main effects of interest, imageability, letter length and word frequency were included as covariates in the response time and error analyses, however their inclusion did not change the pattern of results (see footnote 2).

As in Experiment 1, to measure semantic distance we used the LSA (Landauer et al, 1998). The associatively and categorically related arrays had closer semantic distance than their corresponding unrelated arrays (associatively related vs. unrelated: $t(22) = 3.62, p = .002$; categorically related vs. unrelated: $t(22) = 3.37, p = .002$) but did not show any difference between them (associatively vs. categorically related: $t < 1$). The semantic distance was not different between the two unrelated conditions either ($t < 1$). The LSA results remained the same at a level of $p < .05$ after removing the six words which were not abstract (concreteness ratings > 470).

Two items (i.e., *example, model*) from the categorically related condition had more than 20% error in the translation evaluation task (see Experiment 1 Methods) and given that the two items were presented within the same categorically related array, we excluded the whole array (i.e., *example, model, idol*) from the data analyses.

Design. The experiment consisted of four conditions, each with 12 arrays. Each array consisted of three items. Conditions were identical to those of Experiment 1: associatively related, associatively unrelated, categorically related (near-synonyms), and categorically unrelated. Each condition formed a block in which the 12 arrays were repeated four times, totaling 154 trials per block and 616 in the entire experiment. The order of block presentation was counterbalanced across participants.

Procedure. The procedure for Experiment 2 was the same as Experiment 1, with the one exception that each trial consisted of three rather than four items⁴. Participants heard a Chinese word and were presented with three English words. They selected the English word

that corresponded to the Chinese word. All other components of the procedure were identical to that of Experiment 1. The duration of the experimental period was 40 minutes.

Results and Discussion.

Five participants made more than 20% errors on the language assessment test and were thus excluded from analysis. Eight words were excluded from data analysis due to high concreteness ratings or more than 20% errors in the translation evaluation task. Response times were preprocessed in the same way as in Experiment 1, and 11% of the data points were removed. An overview of the mean response times, error rates, and standard deviations is presented in Tables 2 and 3. The differences in the response times and error rates between the associatively related vs. unrelated and categorically related vs. unrelated conditions for this experiment are presented in Figure 1.

In the error analysis, there was a significant main effect of semantic context ($\chi^2(1) = 91.56, p < .001$), with more errors committed in the categorically related/unrelated conditions (.15) compared to the associatively related/unrelated conditions³ (.09). The main effect of relatedness was also significant ($\chi^2(1) = 104.39, p < .001$) where the error rate in the related conditions (.15) was larger than the unrelated conditions (.09). There was a close to significant interaction between semantic context and relatedness, showing that the error rate difference was bigger in the categorically related vs. unrelated conditions (.09), compared to the difference between the associatively related vs. unrelated conditions (.04) ($\chi^2(1) = 3.2, p = .07$). The simple effect tests showed that error rates were significantly higher in the associatively related condition (.09) vs. the unrelated condition (.05) ($\chi^2(1) = 29.02, p < .001$) and the categorically related (.18) vs. the unrelated condition (.10) ($\chi^2(1) = 92.76, p < .001$).

In the response time analysis, there was a significant main effect of semantic context ($F(1, 61) = 15.27, p < .001$) where the response times in the categorically related/unrelated conditions (1109 ms) were slower than in the associatively related/unrelated conditions (1026 ms). The main effect of relatedness was also significant ($F(1, 10959) = 50.81, p < .001$), where response times in the related conditions (1068 ms) were slower than the unrelated conditions (1049 ms). In contrast to Experiment 1, there was no interaction between semantic context and relatedness ($F's < 1$).

Overall, the results in Experiment 2 with abstract words follow the pattern in Experiment 1 with concrete words, a pattern inconsistent with the different representational frameworks theory (Crutch & Warrington, 2005, 2010), which predicts a larger associative vs. categorical effect for abstract words. As in Experiment 1 with concrete words, for abstract words we observed significant categorical and associative interference effects where the categorical interference effect was larger than the associative effect in the error rate analysis, where this difference just missed statistical significance. Although for response times we found no difference in magnitude between the associative effect and categorical effect, this null effect was likely a result of the speeded aspect of the task. Specifically, with the fast presentation rate, 0 ms inter-stimulus interval, and 2 s deadline in speeded word matching, participants commit more errors on “interference” trials (i.e., categorically related arrays, error rate: 18%). Majority of the errors were timeout errors (i.e., the participants did not respond within the 2 s deadline). This suggests that quite a few long response times were missing in the categorically related condition, leaving no interaction between semantic context and relatedness in the response time analysis. Therefore, the larger categorical vs. associative

effect in Experiment 2 suggests that abstract words rely more on category than association in their representations, inconsistent with the different representational frameworks theory (Crutch & Warrington, 2005, 2010) that predicts a larger associative vs. categorical effect for abstract words. In Experiments 3 and 4, we again test for these effects using different abstract stimuli (Experiment 3) and directly compare concrete and abstract words (Experiment 4).

Experiment 3 – Abstract Words (taxonomic categories)

Although we observed similar results for concrete and abstract words in the first two experiments, it could be argued that the larger synonym effect for abstract words does not truly reflect the category representation of abstract words due to the non-comparable category definitions used for concrete and abstract words (i.e., category for concrete words was defined within a taxonomic structure whereas for abstract words near-synonyms were used). Therefore, in Experiment 3 we re-evaluated whether we could observe a larger categorical vs. associative effect for abstract words when employing the same category definition used for concrete words.

Method

Participants. 32 Chinese-English bilingual participants (Age: 20 years \pm 1.7) took part in this experiment, speaking both English and Chinese more than 10 years (English: 14 years \pm 4.0; Chinese: 18 years \pm 1.8). No participant participated in Experiments 1-2.

Materials, Design, and Procedure. The same number of stimuli (64 abstract English words and their Chinese translations) was used in this experiment as in Experiment 1.

Thirty-two words were used in the categorically related and unrelated condition (24 words from Verheyen et al., 2011). Similarly, 32 words were used in the associatively related and

unrelated conditions (24 words from Experiment 2 and the rest from a subset of materials used by Crutch and Warrington (2007, Experiment 2) and Zhang et al. (2012, Experiment 2)). See Appendix C for Experiment 3 stimuli. In the categorically related condition, there were eight categories: Art forms, Crimes, Diseases, Emotions, Sciences, Virtues, Stages of life, and Time. Similar to Experiment 2, this experiment was also originally run with all 72 words. However, as in Experiment 2, to better control for degree of abstractness before the data analysis we eliminated the words with concreteness ratings above 470 (ratings taken from the MRC psycholinguistic database, Coltheart, 1981). According to this criterion, 10 abstract words (i.e., *painting, cancer, money, bank, market, punch, family, music, malaria, dance*) were too concrete and thus were excluded from the data analyses. Critically, the remaining stimuli had similar concreteness ratings (average: 354) (Coltheart, 1981) as the items in Experiment 2 (concreteness ratings: $t < 1$), and had significantly lower ratings in concreteness than the stimuli in Experiment 1 (concreteness: $t(116) = 20.90; p < .001$).

Following the associative and categorical rating task (see Experiment 1 Methods), the associative ratings for the associatively related word arrays (average: 6.28) were significantly higher than the associatively unrelated word arrays (average: 1.25) ($t(14) = 4.08, p < .001$). Similarly, the categorical ratings for the categorically related arrays (average: 6.33) were significantly higher than the categorically unrelated arrays (average: 1.15) ($t(14) = 31.73, p < .001$). Importantly, the associative ratings for the associatively related arrays (average: 6.28) were significantly higher than the associative ratings for the categorically related arrays (average: 3.88) ($t(14) = 6.37, p < .001$). The categorical ratings for the categorically related

arrays (average: 6.33) were significantly higher than the categorical ratings for the associatively related arrays (average: 2.57) ($t(14) = 10.74, p < .001$).

The two sets of stimuli did not differ in terms of imageability, concreteness, word frequency and word length (Coltheart, 1981) (p 's $> .15$, see Table 1). The categorically and associatively related arrays had closer semantic distance (LSA, Landauer et al, 1998) than their corresponding unrelated arrays (associatively related vs. unrelated: $t(14) = 5.23, p < .001$; categorically related vs. unrelated: $t(14) = 5.14, p < .001$) but did not show any difference between them (associatively vs. categorically related: $t(14) = 1.47, p = .16$). The semantic distance was not different between the two unrelated conditions either ($t < 1$). The LSA semantic distance results remained the same at a level of $p < .05$ after removing the 10 words with high concreteness ratings (see above). Following the translation evaluation task (see Experiment 1 Methods), all materials had lower than 20% error rate in translation and were thus included in the data analysis. The design and procedure in this experiment were identical to Experiment 1.

Results and Discussion

All participants made less than 20% errors on the language assessment test. However, we removed from analysis two participants who had error rates of more than three standard deviations from the mean error rate. Ten words were excluded from data analysis due to high concreteness ratings. Response times were preprocessed in the same way as in Experiments 1-2, and 13% of the data points were removed. An overview of the mean response times, error rates, and standard deviations is presented in Tables 3 and 4. The differences in the response times and error rates between the associatively related vs. unrelated and

Similar to Experiment 1 with concrete words, we observed for abstract words significant categorical and associative interference effects where the categorical interference effect was larger than the associative effect in both the error rate and response time analyses. In this experiment, category for abstract words was defined within a taxonomic structure (Verheyen et al., 2011), as was done for concrete words in Experiment 1 and concrete words elsewhere (e.g., Crutch & Warrington, 2007). The significantly larger categorical effect for abstract words observed in Experiment 3 (and marginally significant in Experiment 2) suggests that the representation of abstract words relies more on category than association.

Importantly, the different representational frameworks theory (Crutch & Warrington, 2005, 2010) assumes that the representations of abstract and concrete concepts differ in terms of category and association. Specifically, the theory predicts that there should be a larger categorical than associative effect for concrete compared to abstract words (i.e., a significant three-way interaction between word type (concrete/abstract), semantic context (categorical/associated) and relatedness (related/unrelated)). When analyzing the data collapsed across Experiment 1 (concrete) and Experiment 3 (abstract) we found no significant three-way interaction (word type, semantic context and relatedness) in either error ($F < 1$) or the response time analyses ($F(1, 25107) = 1.68, p = .20$). Although this suggests that the representations of concrete and abstract words were similar in terms of category and association, the non-significant interaction might be due to the lack of power inherent in a between-participant design. We thus conducted Experiment 4 to address this question using a within-participant design.

Experiment 4 – Abstract and Concrete Words

To provide converging evidence for the similarly larger categorical effects observed for both concrete and abstract words across Experiments 1-3, in Experiment 4 we used a within-participant design to evaluate whether the magnitudes of categorical and associative effects differ between concrete and abstract words. As none of the previous studies with healthy or brain-damaged individuals (e.g., Crutch & Warrington, 2005, 2007, 2010; Hamilton & Martin, 2011; Crutch et al., 2009; Zhang et al., 2012) reported the three-way interactions between word type, semantic context, and relatedness we presented above, they cannot provide an effect size estimate. Hence, to estimate the sample size required to achieve a power of .80 to detect a significant interaction between word type, semantic context, and relatedness with a within-participant design we used the partial eta squared effect sizes (RT and error) from Experiment 3 using ANOVA repeated measures estimates⁵ (G*Power 3, Faul, Erdfelder, Lang, & Buchner, 2007). The effect sizes were .01 for the error analysis and .009 for the response time analysis. Thus, for a within-participant design, to achieve power of .80 to detect a significant three-way interaction, 44 participants were required for the error analysis and 46 for the RT analysis. Consequently in Experiment 4, to provide converging evidence for the similarly larger categorical effects observed for both concrete and abstract words across Experiments 1-3, we used a within-participant design (n=46) to evaluate whether the magnitudes of categorical and associative effects differ between concrete and abstract words.

Method

Participants. 46 Chinese-English bilingual participants (Age: 19 years \pm 1.3) took part in this experiment, speaking both English and Chinese more than 10 years (English: 13 years \pm 3.2; Chinese: 17 years \pm 3.4). No participant participated in Experiments 1-3.

Materials, Design, and Procedure. The majority of the stimuli from Experiment 1 (20 words in the categorically related/unrelated conditions and 29 words in the associatively related/unrelated conditions) and all the stimuli from Experiment 3 were used in this experiment. For the concrete word stimuli in Experiment 4, we replaced 15 words from Experiment 1 stimuli (see Appendix D) in order to control for differences between categorically related/unrelated and associatively related/unrelated conditions in terms of concreteness, imageability, and word frequency (p 's $>$.10, ratings taken from the MRC psycholinguistic database, Coltheart, 1981). We removed 10 abstract words from the data analysis as we did in Experiment 3. As a result, the concrete words had higher concreteness (average: 602; range 532-646) (Coltheart, 1981) compared to the abstract words (average: concreteness: 354, range 261 – 470; $t(116) = 21.44$; $p < .001$). Following the associative and categorical rating task (see Experiment 1 Methods), for concrete words the associative ratings for the associatively related word arrays (average: 6.85) were significantly higher than the associatively unrelated word arrays (average: 1.25, $t(14) = 64.48$, $p < .001$). Similarly, for concrete words the categorical ratings for the categorically related arrays (average: 6.28) were significantly higher than the categorically unrelated arrays (average: 1.07) ($t(14) = 28.34$, $p < .001$). Importantly, the associative ratings for the associatively related arrays (average: 6.85) were significantly higher than the categorically related arrays (average: 4.36) ($t(14) = 5.32$, $p < .001$). The categorical ratings for the categorically related arrays (average:

6.28) were significantly higher than the associatively related arrays (average: 2.56) ($t(14) = 19.60, p < .001$).

The categorically and associatively related arrays had closer semantic distance (LSA, Landauer et al, 1998) than their corresponding unrelated concrete words arrays (associatively related vs. unrelated: $t(14) = 6.99, p < .001$; categorically related vs. unrelated: $t(14) = 4.56, p < .001$) but did not show any difference between them (associatively vs. categorically related: $t < 1$). The semantic distance was not different between the two unrelated conditions either ($t < 1$). We recruited an additional eight English-Chinese bilingual participants who did not participate in Experiments 1-4 (Age: 19 years \pm 1.1; English/Chinese bilinguals for more than 10 years (English: 14 years \pm 4.1; Chinese: 19 years \pm 3.1)), to complete a translation evaluation task for the 15 new concrete words (see Experiment 1 Methods). Because one item (*dress*) in the categorically related condition had more than 20% error, we excluded this item from the data analysis. The design and procedure in this experiment were identical to Experiment 1 except that all participants completed both concrete and abstract word trials, blocked by condition. The order of condition blocks was counterbalanced across participants.

Results and Discussion

Three participants made more than 20% error on the language assessment test. We did not include their data in the data analysis. Due to more than 20% errors for the item (i.e., *dress*) in the translation evaluation task, we excluded it from the data analysis. Response times were discarded from the analyses whenever: (a) an incorrect English word was chosen; (b) there was no response; or (c) response times deviated from a participant's mean by more than three standard deviations resulting in exclusion of 13% of the data. Semantic context and

word type were within-participant and between-item variables. Relatedness was a within-participant and within-item variable. Using linear mixed-effects modeling (e.g., Baayen et al., 2008) for the response time analysis, random factors included participants and items (nested within word type). Fixed factors included in the models were word type (concrete vs. abstract), semantic context (category vs. association), relatedness (related vs. unrelated), the interaction between semantic context and relatedness, the interaction between word type and semantic context, the interaction between word type and relatedness, and the interaction between word type, semantic context, and relatedness. Logistic regression (e.g., Agresti, 2002) was employed for the error analyses in this experiment. Fixed factors were the same as for the linear mixed-effects model. An overview of the mean response times, error rates, and standard deviations is presented in Tables 2 and 3. The differences in the response times and error rates between the associatively related vs. unrelated and categorically related vs. unrelated conditions for this experiment are presented in Figure 1.

In the error analysis, there was a significant main effect of word type ($\chi^2(1) = 155.10, p < .001$), where participants made more errors identifying abstract words (.15) vs. concrete words (.10). The main effect of relatedness was also significant ($\chi^2(1) = 321.67, p < .001$), where participants made more errors in the categorically/associatively related conditions (.15) than the unrelated conditions (.09). There was a significant interaction between semantic context and relatedness ($\chi^2(1) = 5.64, p = .02$) where the error rate difference was bigger in the categorically related vs. unrelated condition (.08), compared to the difference between the associatively related vs. unrelated condition (.06). The simple effect tests showed that error rates were significantly higher in the categorically related (.16) vs. the unrelated condition

(.09) ($\chi^2(1) = 207.24, p < .001$) and the associatively related condition (.15) vs. the unrelated condition (.09) ($\chi^2(1) = 113.05, p < .001$). There was no interaction between word type, semantic context and relatedness, suggesting that abstract and concrete words are represented similarly in terms of category and association ($\chi^2(1) = 1.60, p = .21$).

In the response time analysis, there was a main effect of relatedness $F(1, 35128) = 240.91, p < .001$, where response times in the related conditions (1300 ms) were slower than the unrelated conditions (1252 ms). There was a significant interaction between semantic context and relatedness ($F(1, 35128) = 11.65, p < .001$), where the size of the categorical effect (57 ms) was larger compared to the associative effect (39 ms). The simple effect tests showed that there was a significant difference in the response times between associatively related (1300 ms) and unrelated (1261 ms) items ($F(1, 35128) = 74.19, p < .001$), and a significant difference between categorically related (1299 ms) and unrelated (1242 ms) items ($F(1, 35128) = 177.16, p < .001$). There was no interaction between word type, semantic context and relatedness, suggesting that abstract and concrete words are represented similarly in terms of category and association ($F(1, 35128) = 1.74, p = .19$).

Consistent with Experiments 1-3, using a within-participant design to test whether translating concrete and abstract words is affected by categorical vs. associative relationships, Experiment 4 demonstrated the associative and categorical effects but a larger categorical effect for both concrete and abstract words. This pattern suggests that concrete and abstract concepts are represented by both category and association but rely more on category, a pattern inconsistent with the different representational frameworks theory (Crutch & Warrington, 2005, 2010).

General Discussion

We present the first study with healthy participants to test two predictions from the different representational frameworks theory (Crutch & Warrington, 2005, 2010) that concrete and abstract concepts are represented both by category and association, where concrete concepts are represented primarily by category and abstract concepts by association. We tested these two predictions for concrete words (Experiment 1), and abstract words using the representing criterion (i.e., synonyms) from Crutch and Warrington (2005, 2007, 2010; Experiment 2) and further investigated the category structure for abstract words in a way more similar to that for concrete words (i.e., both taxonomically, e.g., *animals: cat and rabbit; sciences: physics and biology*; Experiment 3), while controlling for semantic distance. Critically, we directly compared abstract and concrete concepts in terms of category and association in Experiment 4. Across four experiments, when matching concrete and abstract Chinese words to the corresponding English words from among multiple categorically (or synonyms) related, associatively related or unrelated distractors, we observed higher error rates and longer response times in the two related conditions compared to the unrelated conditions. Critically, the size of the categorical effect was larger than the associative effect for both concrete and abstract targets and there was no difference between concrete and abstract words. Together, our results are inconsistent with the different representational frameworks theory. Below we discuss how our results fit with previous evidence and two potential theoretical accounts.

That concrete words were more difficult to translate in the context of categorically/ associatively related vs. unrelated words in our experiments is consistent with previous

studies in healthy and aphasic participants. As discussed in the Introduction, seven patients across different studies (A.Z., R.O.M., and F.B.I. in Crutch & Warrington, 2005, 2007, 2010; I.R.Q. in Crutch et al., 2006; UM-103 in Hamilton & Coslett, 2008; H.A. and D.Z. in Hamilton & Martin, 2011) committed more errors in the categorically related vs. the unrelated condition for concrete words, while three of seven (UM-103 in Hamilton & Coslett, 2008; H.A. and D.Z. in Hamilton & Martin, 2011) committed more errors in the associatively related vs. unrelated condition for concrete words. Elsewhere across a variety of explicit and implicit tasks (e.g., free recall, picture-matching, lexical decision), both children and adults categorize concrete objects in terms of category and association (Blanchet, Dunham, & Dunham, 2001; Hutchison, 2003; Lucariello, Kyratzis, & Nelson, 1992; Lin & Murphy, 2001; Mirman & Graziano, 2012; Murphy, 2001; Waxman & Namy, 1997). For example, in a picture-matching task where children were asked to choose the picture that “goes best with” the target picture from an array of categorical and associatively related concrete objects, children selected both categorical and associative choices (e.g., Blanchet, Dunham, & Dunham, 2001; Waxman & Namy, 1997). Interestingly, young children prefer to use associations (e.g., *dog-bone*) to group concrete objects but not categories (e.g., *animal: dog-cat*), whereas older children and adults rely more on categories, a change referred to as an associative-to-categorical shift (e.g., Baldwin, 1989; Waxman & Kosowski, 1990; Waxman & Gelman, 1986; but see Lin & Murphy, 2001). This phenomenon is consistent with the larger categorical effects in Experiments 1 and 4 (concrete words). Moreover, a number of priming studies with healthy participants reported faster response times in a lexical decision and word reading task with concrete words when the target word (e.g., *dog*) was presented

after either a categorically or associatively related prime (e.g., *cat* or *bone*) compared to an unrelated prime (e.g., *table*) (see Hutchison, 2003, for a review). This suggests that once activated, a word spreads activation to categorically and associatively related words, facilitating the recognition of target words. Thus, these studies suggest that category and association are both representing principles for concrete words.

Similarly, the categorical and associative effects in Experiments 2-4 indicate that the representations of abstract concepts also depend on category and association. Unlike concrete concepts, previous research on abstract concepts is restricted to exploring the processing differences between concrete and abstract words (e.g., James, 1975; Kroll & Merves, 1986; Paivio, 1971; Strain, Patterson, & Seidenberg, 1995; see Paivio, 1991, for a review) rather than the representing principles underlying abstract concepts. One exception is two lexical decision priming studies which reported faster response times for the target word (e.g., *quiet*) after a synonym prime (e.g., *calm*) compared to an unrelated prime (e.g., *idea*) for both healthy and brain-damaged participants (Bleasdale, 1987; Tyler, Moss, & Jennings, 1995), results consistent with the categorical (synonym) effect observed in Experiment 2. Additionally, associative and categorical effects were also both observed for abstract concepts in two recent patient studies discussed in the Introduction where three patients made more errors in both categorically (synonym) and associatively related conditions than the unrelated condition in a word-matching task (Hamilton & Coslett, 2008; Hamilton & Martin, 2010). Thus, consistent with the above, whether abstract word categories were synonyms (Experiment 2) or taxonomic categories (Experiment 3 and 4), we found both categorical and associative interference effects for abstract words in healthy participants suggesting that

abstract concepts also rely on category and association.

There are at least two general theoretical accounts for the larger categorical effect we observed for both concrete and abstract words. First, the larger categorical effect in concrete words could be interpreted by the association-to-category shift described above (i.e., young children rely on associations to classify concrete objects while older children and adults rely on taxonomic categories) due to cognitive development (e.g., Baldwin, 1989; Waxman & Kosowski, 1990; Waxman & Gelman, 1986; but see Lin & Murphy, 2001). According to a cognitive development account (Inhelder & Piaget, 1964; Markman, 1989; Vygotsky, 1962; Waxman & Markow, 1995), children first rely on observations (e.g., The *dog* is chewing the *bone*) in daily life to categorize concrete objects and then start forming taxonomic categories once they have developed logical thinking and generalization ability. Due to the fact that objects in the same associative relationship form familiar scenes or events that children encounter in their daily life, young children rely on associations to group objects (e.g., Baldwin, 1989; Waxman & Kosowski, 1990). In comparison to association, learning taxonomic categories requires more logical thinking and generalization. For example, in order to know that “a *dog* is an *animal*” but “a *toy dog* is not” (although they look very similar), children have to clearly understand the critical properties for animals (e.g., *has skin, eats, breathes, can move around*). Thus, it has been argued that children acquire taxonomic categories later than associations (Vygotsky, 1962; Inhelder & Piaget, 1964). Once children acquire taxonomic categories, they start using them to classify objects more than associations, as categories provide a more efficient way to learn new words, identify new objects and communicate with others (e.g., Markman, 1989). For example, once children learned that

“*dogs can eat*”, they quickly identified that “*cats can eat too*”, as dogs and cats are animals and “eating” is the critical property for the animal category. Thus, older children and adults rely more on taxonomic category for concrete concepts. Although this account focuses mainly on concrete concepts, following the same logic, children are likely to apply association and category to abstract concepts. Similar to concrete concepts, the feature overlap (e.g., *school, grade*) between members in the same taxonomic category for abstract concepts (e.g., *math, physics*) might also provide a more efficient way for children to learn abstract concepts. Hence, this account can interpret the larger categorical effect for both concrete and abstract words.

In contrast, an alternative account suggests that feature overlapping is the basis for the larger categorical effect for concrete and abstract words (Collin & Loftus, 1975). According to the spreading-activation theory (Collin & Loftus, 1975), our knowledge is a conceptual network organized in terms of semantic similarity. More specifically, “The more properties two concepts have in common, the more links there are between the two nodes via these properties and the more closely related are the concepts” (p. 411, Collin & Loftus, 1975). Hence, in this network, the concepts (e.g., *car, bus*) within the same taxonomic category (e.g., *vehicle*) are more closely clustered together in comparison to associatively related concepts (e.g., *car, street*), due to more physical attributes or functions shared by the same taxonomic category members (e.g., *cars and buses have engines, wheels, and are used to carry people from place to place*) (see p. 412, Figure 1, Collin & Loftus, 1975). Under this interpretation, in our translation task, once participants heard the Chinese target word, this concept spreads activation to categorically/associatively related concepts but the categorically related

concepts received greater activation compared to those associatively related concepts due to more features shared between the categorically related concepts. Consequently, this organization results in more difficulty in selecting the corresponding English word among the categorically vs. associatively related words. Although the spreading-activation theory focuses mainly on the representations of concrete concepts, following the same logic, the feature overlap between members within the same taxonomic category for abstract words might also lead to a larger categorical effect for abstract words. Hence, the spreading-activation theory accounts for the larger categorical effect for both concrete and abstract words.

In support of a feature overlapping account, several studies demonstrate that the way we process these types of concepts is potentially due to the types of taxonomic features associated with each category (sensory, motor, or affective; e.g., James, 1975; Kroll & Merves, 1986; Paivio, 1971; Strain, Patterson, & Seidenberg, 1995; see Paivio, 1991, for a review). Although taxonomic category is important for concrete and abstract concepts, different semantic features are likely involved in the category structure between the two types of words. For example, Wiemer-Hastings and Xu (2005) systematically compared the features associated with abstract vs. concrete concepts by using a feature generation task where participants produced the features for 18 concrete and 18 abstract concepts. More features expressing participative experiences such as emotion (e.g., *good*) were associated with abstract concepts (e.g., *hope*) whereas more sensory/motor features (e.g., *green color*) were associated with concrete concepts (e.g., *tree*). Furthermore, Kousta et al. (2011) proposed that both concrete and abstract concepts involve experiential (sensory, motor, and

affective) and linguistic information but they differ in terms of the weight of sensory, motor, or affective information. Experiential information is more relevant to taxonomic category structure for both concrete and abstract words, as it relates to physical attributes or properties associated with concepts. Linguistic information refers to verbal associations based on co-occurrence and syntactic information, similar to the association structure used in the different representational frameworks theory (Crutch & Warrington, 2005, 2010). Kousta et al. (2011) further argued that sensory/motor information is more important for concrete concepts whereas affective information is more important for abstract concepts, as affective information accounts for the difference in processing concrete and abstract words. Therefore, although the current results suggest we use both category and association to process concrete and abstract words, future empirical research should further investigate whether abstract and concrete categories differentially rely on sensory, motor, and affective information in terms of their categorical structure, in both children and adults.

We wish to point out that it is important for our and other studies to consider how the categories and relationships between concepts are defined. Our long-term knowledge about the world is generally thought to be divided along fundamental lines however the criteria for where those lines fall have been open for debate for a long time. These distinctions have been most noticed when, as the result of brain-damage, one type of knowledge is impaired and another is left relatively intact. For example, double dissociations in our knowledge are observed between information concerning episodic vs. semantic knowledge (e.g., Butters, Granholm, Salmon, Grant, & Wolfe, 1987; Squire & Zola, 1998; Tulving, 1985), living vs. non-living things (e.g., Warrington & Shallice, 1984; Caramazza & Shelton, 1998), abstract

vs. concrete concepts (e.g., Warrington, 1975; Warrington & Shallice, 1984), or how concepts are related to each other, for example categorical vs. associative relationships (e.g., Crutch & Warrington, 2003, 2007, 2010).

The exact definitions of these important theoretical constructs continue to be debated. For example, concreteness and imageability ratings have been used interchangeably to define concrete and abstract words (e.g., Binder, Westbury, McKiernan, Possing, & Medler, 2005; Fliessbach, Weis, Klaver, Elger, & Weber, 2006; Richardson, 2003). However, concreteness and imageability tap into different aspects of concepts (Kousta, Vigliocco, Vinson, Andrews, & Campo, 2011). Specifically, the frequency distribution of concreteness ratings is bimodal, with two distinct modes for concrete and abstract words (see also Nelson & Schreiber, 1992), whereas the distribution of imageability ratings is unimodal. In other words, concreteness ratings better capture a fundamental difference between concrete and abstract words, whereas imageability ratings measure graded sensory (primarily visual) properties associated with concrete and abstract words (Kousta et al., 2011). As such, we chose to use concreteness ratings to define the concrete and abstract words in our study, as concreteness ratings are a more appropriate way to define concrete vs. abstract words.

Similarly, the definitions of categorical and associative relationships are also typically under-specified. Different terms are used to describe semantic relationships where it is unclear whether these terms refer to the same relationship. For example, to refer to the notion that two concepts typically co-occur either in language or in real-life situations, both “associative” and “thematic” are used interchangeably. However, an associative relationship refers to linguistic co-occurrence in most studies where association is operationally defined in

terms of free association probabilities (i.e., the likelihood of producing a given target word in response to a specific cue word; e.g., see Hutchison, 2003 for a review). Specifically, there are a number of ways in which two words can be associated, such as their co-occurrence in conjoined noun phrases (e.g., *foot-ball*), as category coordinates (e.g., *dog-cat*), and via a thematic relationship (e.g., *dog-bone*). In contrast, thematically related concepts perform complementary roles in the same scenario or event (e.g., Estes, Golonka, & Jones, 2011; Lin & Murphy, 2001). According to this definition, many associated concepts are not thematically related. Further complicating how our knowledge is thought to be divided, the thematic/associative and taxonomic/categorical relationships are not mutually exclusive (see Estes, Golonka, & Jones, 2011). For example, dog and cat can form a familiar scene (e.g., the dog is chasing a cat, a thematic/associative relationship) and the same category (i.e., animals, a categorical/taxonomic relationship). Given that sometimes concepts can be both thematically and taxonomically related, in our study, we collected participants' ratings for categorical and associative strength respectively and used identical instructions for both concrete and abstract words to provide better measurements of categorical and associative strength. Here, we chose to operationalize associative and categorical relationships based on the criteria used in several recent studies (e.g., Mirman & Graziano, 2012; Schwartz et al., 2011; Zhang et al., 2012). We acknowledge that the boundaries defining these terms are not fixed and that there may be other ways to operationalize these concepts. It is an important question for future research to understand whether our results concerning these types of representations generalize to differently described distinctions between types of knowledge.

In summary, by improving upon previous study designs and employing a new

comprehension task, this study provides novel evidence in healthy participants regarding the role of semantic relations in the representations of concrete and abstract words. We tested predictions not previously examined in healthy people from the different representational frameworks theory of the representations of concrete and abstract concepts (Crutch & Warrington, 2005, 2010). Inconsistent with this theory, our results suggest that abstract and concrete concepts are represented similarly in terms of category and association.

References:

Baldwin, D. (1989). Priorities in Children's Expectations about Object Label Reference:

Form over Color. *Child Development*, 60(6), 1291–1306.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher &

R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 129–163). Cambridge: Cambridge University Press.

Blanchet, N., Dunham, P. J., & Dunham, F. (2001). Differences in preschool children's

conceptual strategies when thinking about animate entities and artifacts. *Developmental Psychology*, 37, 791–800.

Bleasdale, F. A. (1987). Concreteness-dependent associative priming: Separate lexical

organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(4), 582-594.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word

translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, 48, 468–488.

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005).

Distinct brain systems for processing concrete and abstract words. *Journal of Cognitive Neuroscience*, 17, 905–917.

Butters, N., & Granholm, E., & Salmon, D. (1987). Episodic and semantic memory: a

comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology*, 9, 479 – 497.

- Capitani, E., Laiacona, M., Mahon, B., & Caramazza, A. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology, 20*, 213–61.
- Caramazza, A., & Costa, A. (2000). The semantic interference effect in the picture–word interference paradigm: Does the response set matter? *Cognition, 75*, B51–B64.
- Caramazza, A., & Costa, A. (2001). Set size and repetition in the picture–word interference paradigm: Implications for models of naming. *Cognition, 80*, 291–298.
- Caramazza, A., & Shelton, J. R. (1998). Domain specific knowledge systems in the brain: the animate-inanimate distinction. *Journal of Cognitive Neuroscience, 10*, 1–34.
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 33*, 497 – 505.
- Crutch, S. J. (2006). Qualitatively different semantic representations for abstract and concrete words: Further evidence from semantic reading errors of deep dyslexic patients. *Neurocase, 12*, 91–97.
- Crutch, S. J., Connell, S., & Warrington, E. K. (2009). The different representational frameworks underpinning abstract and concrete knowledge: Evidence from odd-one-out judgements. *Quarterly Journal of Experimental Psychology, 62*, 1377–1390.
- Crutch, S. J., & Jackson, E. C. (2011). Contrasting graded effects of semantic similarity and association across the concreteness spectrum. *The Quarterly Journal of Experimental Psychology, 64(7)*, 1388–1408.

- Crutch, S. J., Ridha, B. H., & Warrington, E. K. (2006). The different frameworks underlying abstract and concrete knowledge: Evidence from a bilingual patient with a semantic refractory access dysphasia. *Neurocase, 12*, 151–163.
- Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain, 128*, 615–627.
- Crutch, S. J., & Warrington, E. K. (2010). The differential dependence of abstract and concrete words upon associative and similarity-based information: Complementary semantic interference and facilitation effects. *Cognitive Neuropsychology, 27*, 46-71.
- De Groot, A., & Poot, R. (1997). Word Translation at Three Levels of Proficiency in a Second Language: The Ubiquitous Involvement of Conceptual Memory. *Language learning, 47*, 215 – 264.
- Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. *Psychology of Learning and Motivation, 54*, 249–294.
- Estes, Z., & Jones, L. L. (2009). Integrative priming occurs rapidly and uncontrollably during lexical processing. *Journal of Experimental Psychology: General, 138*, 112–130.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.
- Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage, 32*, 1413–1421.
- Forster, K. I., & Forster, J. C. (2003). A Windows display program with millisecond accuracy.

- Behavior Research Methods Instruments & Computers*, 35, 116-124.
- Gilhooly, K.J. and Logie, R.H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behaviour Research Methods and Instrumentation*, 12, 395-427.
- Hamilton, A. C., & Coslett, H. B. (2008). Refractory access disorders and the organization of concrete and abstract semantics: Do they differ? *Neurocase*, 14, 131–140.
- Hamilton, A. C., & Martin, C. R. (2010). Inferring Semantic Organization from Refractory Access Dysphasia – Further Replication in the Domains of Geography and Proper Nouns but not Concrete and Abstract Concepts. *Cognitive Neuropsychology*, 27, 614-635.
- Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, 10, 785–813.
- Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child*. London: Routledge & Kegan Paul.
- James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, 104, 130–136.
- Kalénine, S., Peyrin, C., Pichat, C., Segebarth, C., Bonthoux, F., & Baciú, M. (2009). The sensory–motor specificity of taxonomic and thematic conceptual relations: A behavioral and fMRI study. *Neuroimage*, 44, 1152–1162.
- Kousta, S., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2010). The representation of abstract words: Why emotions matter. *Journal of Experimental Psychology: General*, 140, 14–34.
- Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal*

- of Experimental Psychology: Learning, Memory, and Cognition*, 12, 92–107.
- Kroll, J. K., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174
- Kucera and Francis, W.N. (1967). *Computational Analysis of Present-Day American English*. Providence: Brown University Press.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analyses. *Discourse Processes*, 25, 259-284.
- Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology. General*, 130, 3–28.
- Lucariello, J., Kyratzis, A., & Nelson, K. (1992). Taxonomic knowledge: What kind and when? *Child Development*, 63, 978–998.
- Mahon, B. Z., and Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annual Review of Psychology*, 60, 1–15.
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 503–535.
- Markman, E. M.(1989). *Categorization and naming in children: Problems of induction*. Cambridge: MIT Press.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.

- McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 558–572.
- Mirman, D. & Graziano, K.M. (2012). Individual differences in the strength of taxonomic versus thematic relations. *Journal of Experimental Psychology: General*.
- Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 863–883.
- Murphy, G. L. (2001). Causes of taxonomic sorting by adults: A test of the thematic-to-taxonomic shift. *Psychonomic Bulletin & Review*, 8, 834–839.
- Navarrete, E., & Costa, A. (2009). The distractor picture paradox in speech production: evidence from the word translation task. *Journal of Psycholinguistic Research*, 38, 527–547.
- Pavio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255–287.
- Paivio, A., Yuille, J.C. and Madigan, S.A. (1968). Concreteness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement*, 76 (3, part 2).
- Richardson, J. (2003). Dual coding versus relational processing in memory for concrete and abstract words. *European Journal of Cognitive Psychology*, 15, 481–509.
- Sachs, O., Weis, S., Krings, T., Huber, W., & Kircher, T. (2008). Categorical and thematic knowledge representation in the brain: Neural correlates of taxonomic

- and thematic conceptual relations. *Neuropsychologia*, *46*, 409–418.
- Sachs, O., Weis, S., Zellagui, N., Sass, K., Huber, W., Zvyagintsev, M., Mathiak, K., & Kircher, T. (2011). How Different Types of Conceptual Relations Modulate Brain Activation during Semantic Priming. *Journal of Cognitive Neuroscience*, *23*, 1263–1273.
- Schwartz, M.F., Kimberg, D.Y., Walker, G.M., Brecher, A., Faseyitan, O., Dell, G.S., Mirman, D., and Coslett, H.B. (2011). A neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences*, *108*, 8520 – 8524.
- Squire, L. R., & Zola-Morgan, S. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*, *8*, 205 – 211.
- Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140–1154.
- Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*, 385–398.
- Tyler, L. K., Moss, H. E., & Jennings, F. (1995). Abstract word deficits in aphasia: Evidence from semantic priming. *Neuropsychology*, *9*(3), 354-363.
- Verheyen, S., Stukken, L., De Deyne, S., Dry, M. J., & Storms, G. (2011). The generalized polymorphous concept account of graded structure in abstract categories. *Memory & Cognition*, *39*, 1117-1132.
- Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. J. M. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*, B61–B69.

- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.
- Warrington, E.K, & Shallice, T. (1984). Category specific semantic impairment. *Brain* 107, 829–54.
- Warrington, E.K. (1975) The selective impairment of semantic memory. *The Quarterly Journal of Experimental Psychology*, 27, 635–57.
- Waxman, S., & Gelman, R.(1986). Preschoolers’ use of superordinate relations in classification and language. *Cognitive Development*, 1, 139-156.
- Waxman, S., & Kosowski, T. (1990). Nouns Mark Category Relations: Toddlers’ and Preschoolers’ Word-Learning Biases. *Child Development*, 61(5), 1461–1473.
- Waxman, S., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology*, 29, 257–302.
- Waxman, S. R., & Namy, L. L. (1997). Challenging the notion of a thematic preference in young children. *Developmental Psychology*, 33, 555–567.
- Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, 29, 719–736.
- Zhang, X., Han, Z., & Bi, Y. (2012). Are abstract and concrete concepts organized differently? Evidence from the blocked translation paradigm. *Applied Psycholinguistics*.

Acknowledgements

We thank Audrey Chao, Megan Kirchgessner, Ibrahim Khan, and Elzia Broussard for collecting the data. We also thank Tao Wei for her helpful suggestions. This study was presented at the 54th Annual Meeting of the Psychonomic Society (2013), Toronto, Ontario, Canada.

Footnotes:

1. There is the possibility that repeating the distractor set may influence participant strategy. Here, if participants were able to predict the correct answer after having excluded the other distractors in previous trials, then response times or error rates would no longer reflect access to semantics (and translation), and instead reflect simply guessing. If participants adopted this strategy, because participants could anticipate response options after the first trial, whether distractors were semantically related or unrelated should have no impact on error rates/RTs and thus any difference between related and unrelated conditions should disappear. However, as will be demonstrated across all four experiments, there were significantly higher error rates and longer RTs in the related vs. unrelated conditions (p 's $< .05$). Thus, the inclusion of repetitions in the experimental design did not impact the semantic effects of interest and allowed us to use the design similar to all previously published papers concerning this issue.
2. To rule out the impact on our main factors of variables that affect the speed of word reading, we also constructed the linear mixed-effects and logistics models using word imageability, word length and word frequency as covariates in all four experiments and observed the same results at a level of $p < .05$.
3. The main effect of semantic context in error rate and response times could potentially be due to unmatched imageability (ratings taken from the MRC psycholinguistic database, Coltheart, 1981). Previous studies show that lower imageability leads to longer response times in word translation tasks (e.g., De Groot & Annette, 1997). The

words in the categorically related/unrelated conditions had higher imageability than the associatively related/unrelated conditions in Experiment 1 (see Table 1), resulting in lower error rates and faster response times in the categorically related/unrelated conditions. Similarly, in Experiment 2, the unmatched imageability ratings between the categorically related/unrelated and the associatively related/unrelated condition may account for the main effect of the semantic context.

4. To ensure distinct Chinese translations for groups of abstract synonyms, we included only three synonyms per group.
5. We chose the program G*Power (Faul et al., 2007) to provide sample size estimates as it is freely and widely available, and is heavily cited across experimental disciplines. However, it has not yet been modified to use parameter estimates from linear mixed effects modeling or logistic regression. Thus, we used values from ANOVA repeated measures to provide sample size estimates in G*Power, as the ANOVA error and response time analyses for all four experiments produced similar results as we found in the linear mixed effects modeling and logistic regressions at a level of $p < .05$.

Tables:

Table 1. Average concreteness (Paivio, Yuille & Madigan 1968), imageability (Gilhooly and Logie, 1980), word frequency (Kucera & Francis, 1967), and word length for categorical and associative stimuli in Experiments 1-4 (after controlling for concreteness ratings resulting in six words removed in Experiment 2 and 10 words removed in Experiments 3 and 4). Exp = Experiment.

		Imageability	Concreteness	Word Frequency	Word length
Exp1(concrete)	category	613	617	29	6
	association	590	596	44	5
Exp2(abstract)	category	424	350	63	5
	association	474	375	83	6
Exp3(abstract) ^a	category	460	356	62	7
	association	463	351	78	7
Exp4(concrete)	category	603	609	45	5
	association	587	595	44	5

^aThe abstract words in Experiment 4 were the same as those as in Experiment 3.

Table 2. Associative and categorical ratings in Experiment 1-4. Exp = Experiment; Assoc.Rel = Associatively related; Assoc.Unrel = Associative unrelated; Categ.Rel = Categorically related; Categ.Unrel = Categorically unrelated.

Experiments	Condition	Associative Rating	Categorical Rating
Exp 1 (concrete)	Assoc.Rel	6.28	2.57
	Assoc.Unrel	1.25	1.18
	Categ.Rel	3.88	6.33
	Categ.Unrel	1.23	1.15
Exp 2 (abstract)	Assoc.Rel	6.07	3.58
	Assoc.Unrel	1.90	1.07
	Categ.Rel	4.69	5.33
	Categ.Unrel	2.24	1.45
Exp 3 (abstract) ^a	Assoc.Rel	6.28	3.79
	Assoc.Unrel	2.45	1.39
	Categ.Rel	2.89	6.33
	Categ.Unrel	1.88	1.15
Exp 4 (concrete)	Assoc.Rel	6.85	2.56
	Assoc.Unrel	1.25	1.20
	Categ.Rel	4.36	6.28
	Categ.Unrel	1.31	1.07

^aThe abstract words in Experiment 4 were the same as those as in Experiment 3.

Table 3. Experiments 1-4: error rate percentages and standard deviations (in parentheses).

	Assoc.Rel	Assoc.Unrel	Diff ^a	Categ.Rel	Categ.Unrel	Diff ^b
Exp1 (concrete)	.13 (.09)	.10 (.07)	.03 (.06)*	.12 (.08)	.06 (.04)	.06 (.07)*
Exp2 (abstract)	.09 (.05)	.05 (.05)	.04 (.05)*	.18 (.10)	.10 (.06)	.08 (.07)*
Exp3 (abstract)	.17 (.08)	.11 (.06)	.06 (.07)*	.20 (.11)	.10 (.08)	.10 (.08)*
Exp4 (concrete)	.12 (.09)	.08 (.05)	.04 (.05)*	.13 (.08)	.07 (.07)	.06 (.05)*
Exp4 (abstract)	.17 (.11)	.11 (.08)	.06 (.07)*	.19 (.12)	.11 (.11)	.08 (.07)*

* $p < .05$. ^a Difference between associatively related and unrelated conditions.

^b Difference between categorically related and unrelated conditions.

Table 4. Experiments 1-4: Response time averages (ms) and standard deviations (in parentheses).

	Assoc.Rel	Assoc.Unrel	Diff ^a	Categ.Rel	Categ.Unrel	Diff ^b
Exp1 (concrete)	1352 (110)	1322 (91)	30 (52)*	1292 (92)	1218 (97)	75 (50)*
Exp2 (abstract)	1048 (125)	1010 (142)	38 (78)*	1128 (137)	1095 (132)	33 (68)*
Exp3 (abstract)	1342 (64)	1305 (82)	37 (67)*	1369 (82)	1307 (89)	62 (68)*
Exp4 (concrete)	1316 (100)	1280 (105)	36 (53)*	1292 (83)	1228 (95)	64 (51)*
Exp4 (abstract)	1284 (111)	1242 (115)	42 (61)*	1306 (122)	1257 (130)	49 (65)*

* $p < .05$. ^a Difference between associatively related and unrelated conditions.

^b Difference between categorically related and unrelated conditions.

Figures:

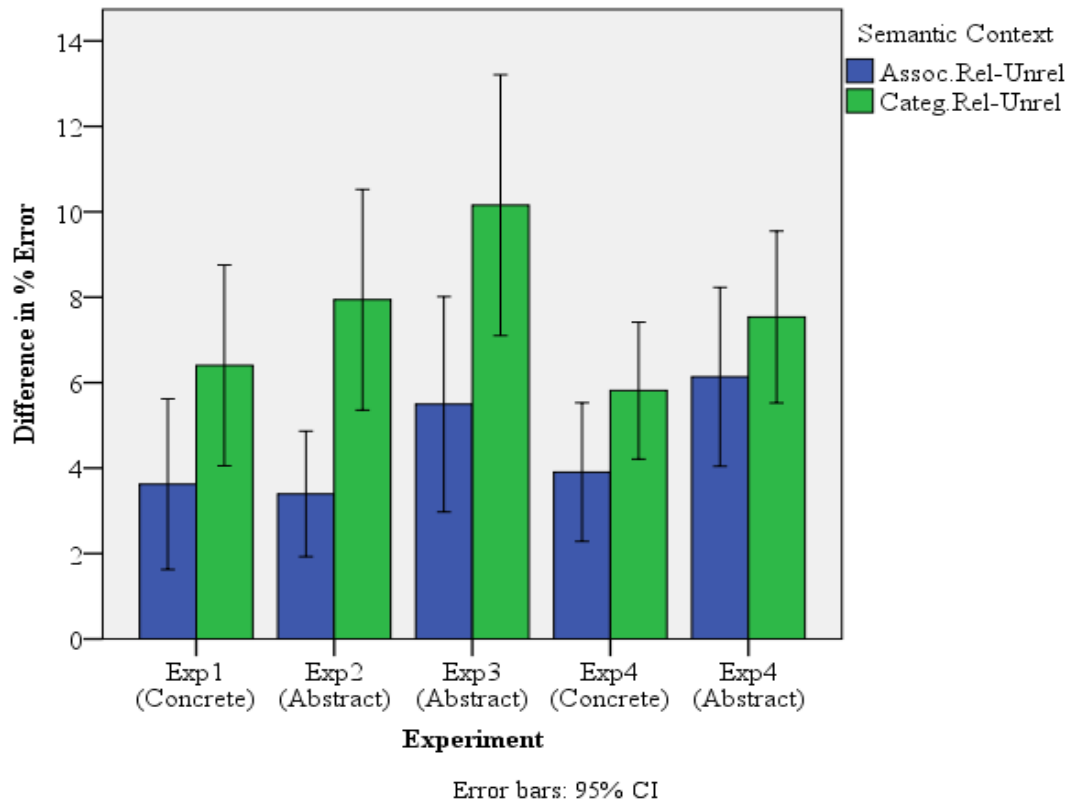


Figure 1a. Error rate differences in the associatively related vs. unrelated (Assoc.Rel-Unrel) and the categorically related vs. unrelated conditions (Categ.Rel-Unrel) in Experiments 1-4. Error bars indicate 95% confidence intervals.

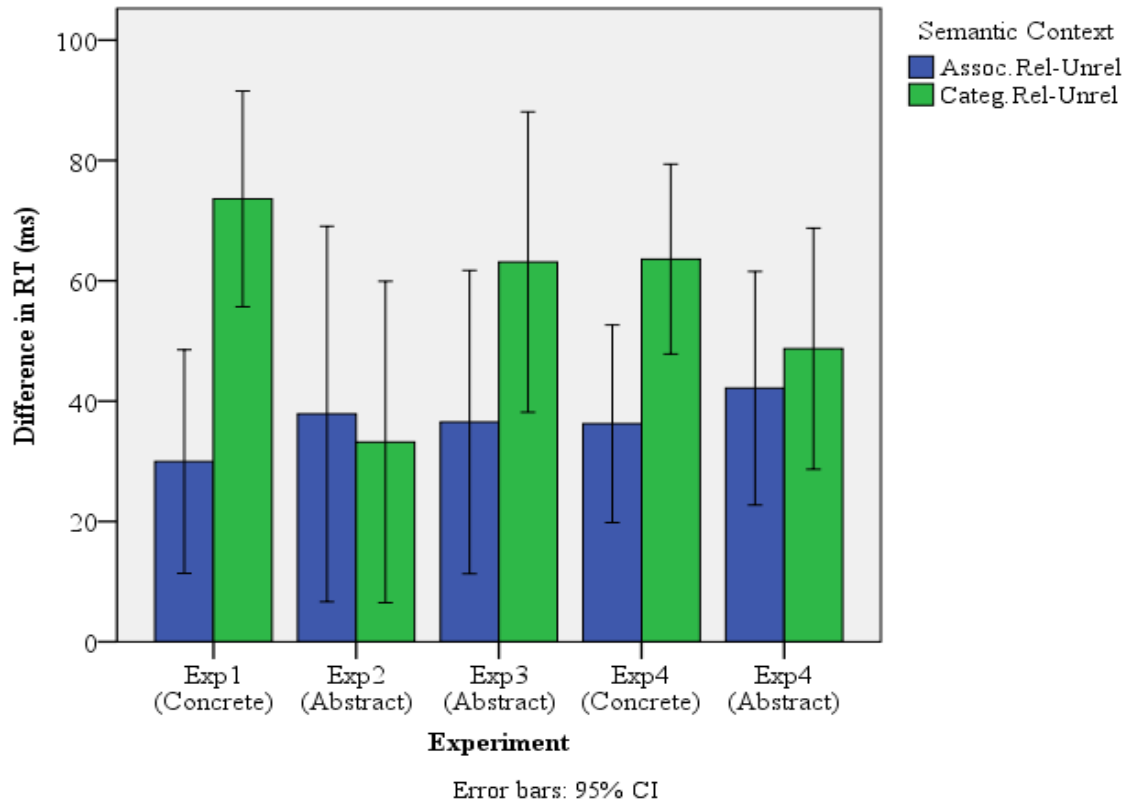


Figure 1b. Response time (RT) differences in the associatively related vs. unrelated (Assoc.Rel-Unrel) and the categorically related vs. unrelated conditions (Categ.Rel-Unrel) in Experiment 1-4. Error bars indicate 95% confidence intervals.

Appendix A: Experiment 1 (concrete words) stimuli.

Concrete words (Associatively Related)

Set 1		Set 2		Set 3		Set 4	
farmer	农民	oven	烤箱	sailor	水手	monkey	猴子
cow	奶牛	apron	围裙	boat	小船	cage	铁笼
tractor	拖拉机	kitchen	厨房	anchor	锚	banana	香蕉
barn	谷仓	soup	汤	ocean	大海	zoo	动物园
Set 5		Set 6		Set 7		Set 8	
tent	帐篷	plane	飞机	picture	图画	building	房屋
camp	野营	airport	机场	brush	画笔	architect	建筑师
woods	森林	ticket	机票	canvas	帆布	cement	水泥
fire	篝火	suitcase	手提箱	frame	画框	brick	砖

Concrete words (Associatively Unrelated)

Set 1		Set 2		Set 3		Set 4	
camp	野营	cow	奶牛	plane	飞机	farmer	农民
oven	烤箱	kitchen	厨房	apron	围裙	soup	汤
anchor	锚	sailor	水手	ocean	大海	boat	小船
monkey	猴子	brick	砖	banana	香蕉	cage	铁笼
Set 5		Set 6		Set 7		Set 8	
tent	帐篷	fire	篝火	woods	森林	tractor	拖拉机
barn	谷仓	airport	机场	ticket	机票	suitcase	手提箱

picture	图画	brush	画笔	canvas	帆布	frame	画框
architect	建筑师	cement	水泥	building	房屋	zoo	动物园

Concrete words (Categorically Related)

Set 1		Set 2		Set 3		Set 4	
skirt	裙子	ship	船	eggplant	茄子	rabbit	兔子
jacket	夹克	car	汽车	pepper	青椒	cat	猫
pants	裤子	bicycle	自行车	onion	洋葱	dog	狗
blouse	衬衫	train	火车	potato	土豆	squirrel	松鼠
Set 5		Set 6		Set 7		Set 8	
cookie	饼干	grapes	葡萄	pigeon	鸽子	violin	小提琴
cake	蛋糕	apple	苹果	duck	鸭子	piano	钢琴
bread	面包	watermelon	西瓜	eagle	老鹰	flute	笛子
sandwich	三明治	strawberry	草莓	goose	鹅	drum	鼓

Concrete words (Categorically Unrelated)

Set 1		Set 2		Set 3		Set 4	
pants	裤子	skirt	裙子	jacket	夹克	cookie	饼干
ship	船	bicycle	自行车	car	汽车	train	火车
eggplant	茄子	pepper	青椒	onion	洋葱	potato	土豆
rabbit	兔子	cat	猫	dog	狗	squirrel	松鼠
Set 5		Set 6		Set 7		Set 8	

blouse	衬衫	cake	蛋糕	sandwich	三明治	bread	面包
grapes	葡萄	apple	苹果	watermelon	西瓜	strawberry	草莓
pigeon	鸽子	duck	鸭子	eagle	老鹰	goose	鹅
violin	小提琴	piano	钢琴	flute	笛子	drum	鼓

Appendix B: Experiment 2 (abstract words) stimuli

Abstract words (Associatively Related)

Set 1		Set 2		Set 3		Set 4	
math	数学	ache	疼痛	network	网络	poverty	贫困
examination	考试	disease	疾病	software	软件	famine	饥饿
result	结果	operation	手术	virus	病毒	refugee	难民
Set 5		Set 6		Set 7		Set 8	
quality	质量	love	爱情	insurance	保险	money	货币
trade	贸易	marriage	婚姻	life	生命	bank	银行
market	市场	family	家庭	safety	安全	savings	储蓄
Set 9		Set 10		Set 11		Set 12	
exercise	锻炼	comedy	喜剧	fight	打架	travel	旅行
healthy	健康	joke	玩笑	violent	暴力	abroad	出国
jogging	慢跑	laugh	大笑	struggle	挣扎	holiday	假期

Abstract words (Associatively Unrelated)

Set 1		Set 2		Set 3		Set 4	
holiday	假期	safety	安全	marriage	婚姻	poverty	贫困
quality	质量	market	市场	software	软件	virus	病毒
ache	疼痛	operation	手术	life	生命	exercise	锻炼
Set 5		Set 6		Set 7		Set 8	
savings	储蓄	love	爱情	family	家庭	fight	打架

trade	贸易	network	网络	result	结果	math	数学
disease	疾病	insurance	保险	abroad	出国	healthy	健康
Set 9		Set 10		Set 11		Set 12	
refugee	难民	violent	暴力	famine	饥饿	struggle	挣扎
examination	考试	joke	玩笑	comedy	喜剧	laugh	大笑
jogging	慢跑	money	货币	travel	旅行	bank	银行

Abstract words(Categorically Related)

Set 1		Set 2		Set 3		Set 4	
anxiety	焦虑	stock	股票	hope	希望	geography	地理
nerves	紧张	fund	基金	ideal	理想	space	空间
fear	害怕	wealth	财产	dream	梦想	location	位置
Set 5		Set 6		Set 7		Set 8	
trick	诡计	signal	信号	example	榜样	idea	想法
steal	偷窃	hint	暗示	model	模范	attitude	态度
deceit	谎言	clue	线索	idol	偶像	viewpoint	观点
Set 9		Set 10		Set 11		Set 12	
ability	能力	taste	品尝	boil	煮沸	clean	干净
skill	技巧	chew	咀嚼	heat	加热	pure	纯净
talent	才干	bite	咬	fry	油炸	fresh	新鲜

Abstract words (Categorically Unrelated)

Set 1		Set 2		Set 3		Set 4	
anxiety	焦虑	fear	害怕	location	位置	hint	暗示
stock	股票	wealth	财产	talent	才干	model	模范
hope	希望	skill	技巧	deceit	谎言	attitude	态度
Set 5		Set 6		Set 7		Set 8	
nerves	紧张	direction	方向	signal	信号	clue	线索
fund	基金	ability	能力	trick	诡计	idol	偶像
ideal	理想	example	榜样	idea	想法	viewpoint	观点
Set 9		Set 10		Set 11		Set 12	
boil	煮沸	space	空间	heat	加热	fry	油炸
clean	干净	fresh	新鲜	pure	纯净	dream	梦想
taste	品尝	bite	咬	chew	咀嚼	steal	偷窃

Appendix C: Experiment 3 (abstract words) stimuli

Abstract words (Associatively Related)

Set 1		Set 2		Set 3		Set 4	
exercise	锻炼	love	爱情	comedy	喜剧	insurance	保险
healthy	健康	marriage	婚姻	joke	玩笑	life	生命
jogging	慢跑	family	家庭	laugh	大笑	safety	安全
muscular	强壮	happiness	幸福	funny	有趣	protection	保护
Set 5		Set 6		Set 7		Set 8	
fight	打架	poverty	贫困	money	货币	quality	质量
violent	暴力	famine	饥饿	bank	银行	trade	贸易
struggle	挣扎	refugee	难民	savings	储蓄	market	市场
punch	拳击	drought	干旱	revenue	税收	merchandise	商品

Abstract words (Associatively Unrelated)

Set 1		Set 2		Set 3		Set 4	
exercise	锻炼	healthy	健康	jogging	慢跑	muscular	强壮
love	爱情	marriage	婚姻	family	家庭	happiness	幸福
comedy	喜剧	joke	玩笑	laugh	大笑	funny	有趣
insurance	保险	life	生命	safety	安全	protection	保护
Set 5		Set 6		Set 7		Set 8	
fight	打架	violent	暴力	struggle	挣扎	punch	拳击
poverty	贫困	famine	饥饿	refugee	难民	drought	干旱

money	货币	bank	银行	savings	储蓄	revenue	税收
quality	质量	trade	贸易	market	市场	merchandise	商品

Abstract words (Categorically Related)

Set 1		Set 2		Set 3		Set 4	
painting	绘画	murder	谋杀	influenza	流感	joy	喜悦
music	音乐	rape	强奸	chickenpox	水痘	grief	悲伤
dance	跳舞	fraud	欺诈	cancer	癌症	fear	害怕
sculpture	雕刻	burglary	盗窃	malaria	疟疾	rage	愤怒
Set 5		Set 6		Set 7		Set 8	
chemistry	化学	honesty	诚实	infancy	婴儿期	day	日子
biology	生物	loyalty	忠诚	childhood	儿童期	century	世纪
physics	物理	compassion	怜悯	adolescence	青春期	month	月份
math	数学	prudence	严谨	adulthood	成人期	decade	十年

Abstract words (Categorically Unrelated)

Set 1		Set 2		Set 3		Set 4	
painting	绘画	chemistry	化学	music	音乐	biology	生物
murder	谋杀	honesty	诚实	rape	强奸	loyalty	忠诚
influenza	流感	infancy	婴儿期	chickenpox	水痘	childhood	儿童期
joy	喜悦	day	日子	grief	悲伤	century	世纪
Set 5		Set 6		Set 7		Set 8	

dance	跳舞	physics	物理	sculpture	雕刻	math	数学
fraud	欺诈	compassion	怜悯	burglary	盗窃	prudence	严谨
cancer	癌症	adolescence	青春期	malaria	疟疾	adulthood	成人期
fear	害怕	month	月份	rage	愤怒	decade	十年

Appendix D: concrete words from Experiment 4 (abstract words were the same as Experiment

3)

Concrete (categorically Related)

Set 1		Set 2		Set 3		Set 4	
skirt	裙子	ship	船	celery	芹菜	chair	椅子
jacket	夹克	truck	卡车	pepper	青椒	table	桌子
pants	裤子	bicycle	自行车	onion	洋葱	couch	沙发
dress	连衣裙	train	火车	potato	土豆	shelf	书架
Set 5		Set 6		Set 7		Set 8	
beer	啤酒	cherry	樱桃	pigeon	鸽子	violin	小提琴
wine	红酒	peach	桃子	duck	鸭子	piano	钢琴
water	水	watermelon	西瓜	eagle	老鹰	flute	笛子
juice	果汁	strawberry	草莓	goose	鹅	drum	鼓

Concrete (categorically Unrelated)

Set 1		Set 2		Set 3		Set 4	
pants	裤子	skirt	裙子	jacket	夹克	beer	啤酒
ship	船	bicycle	自行车	truck	卡车	train	火车
celery	芹菜	pepper	青椒	onion	洋葱	potato	土豆
chair	椅子	table	桌子	couch	沙发	shelf	书架
Set 5		Set 6		Set 7		Set 8	
dress	连衣裙	wine	红酒	juice	果汁	water	水

cherry	樱桃	peach	桃子	watermelon	西瓜	strawberry	草莓
pigeon	鸽子	duck	鸭子	eagle	老鹰	goose	鹅
violin	小提琴	piano	钢琴	flute	笛子	drum	鼓

Concrete (Associatively Related)

Set 1		Set 2		Set 3		Set 4	
farmer	农民	oven	烤箱	sailor	水手	monkey	猴子
cow	奶牛	apron	围裙	boat	小船	cage	铁笼
tractor	拖拉机	kitchen	厨房	anchor	锚	banana	香蕉
barn	谷仓	soup	汤	ocean	大海	zoo	动物园
Set 5		Set 6		Set 7		Set 8	
tent	帐篷	plane	飞机	picture	图画	building	房屋
tree	树	airport	机场	brush	画笔	architect	建筑师
wood	木头	ticket	机票	canvas	帆布	cement	水泥
fire	篝火	suitcase	手提箱	frame	画框	brick	砖
