Applied Psycholinguistics, page 1 of 19, 2013
doi:10.1017/S0142716412000677

Adapting the Pyramids and Palm Trees Test and the Kissing and Dancing Test and developing other semantic tests for the Chinese population

QIHAO GUO
Fudan University

CHENXI HE and XIAOLIANG WEN
Beijing Normal University

LUPING SONG
China Rehabilitation Research Center

ZAIZHU HAN and YANCHAO BI
Beijing Normal University

Received: April 14, 2011 Accepted for publication: August 23, 2012

ADDRESS FOR CORRESPONDENCE
Yanchao Bi, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, People's Republic of China. E-mail: ybi@bnu.edu.cn

ABSTRACT
The semantic system is a core component underlying many cognitive functions, and its deterioration can lead to various behavioral deficits. The Pyramids and Palm Trees Test (Howard & Patterson, 1992) and the Kissing and Dancing Test (Bak & Hodges, 2003) are among the most widely used semantic assessments, and they have been adapted into many languages and for many populations. We adapted these tests to the Mainland Chinese population by adopting culturally appropriate items and collecting normative data in healthy Chinese participants. In addition, accumulating research has shown that semantic representations have multidimensional structures that include various types of knowledge, and in the Pyramids and Palm Trees Test mainly evaluates associative relationships. Therefore, we developed additional tests to examine three other aspects of semantic knowledge: taxonomic, functional, and manipulative. We found significant correlations among all tests in healthy participants. Moreover, the level of education and age affected performance on the tests of associative relationships, taxonomic relationships, and manipulation similarity.

Semantic memory is a long-term memory system for meanings, understandings, and other concept-based knowledge unrelated to specific experiences (Tulving, 1972). It is involved in various cognitive processes, including language processing, object recognition, and object use, and semantic memory impairment leads to a wide range of cognitive deficits (e.g., Buxbaum, Schwartz, & Carew, 1997; Patterson, Nestor, & Rogers, 2007; Warrington, 1975).

The Pyramids and Palm Trees Test (PPT; Howard & Patterson, 1992) is commonly used in clinical and research settings to evaluate the integrity of semantic knowledge. In this test, participants are simultaneously shown three stimuli and asked to identify which of the two items on the bottom (e.g., a palm tree and a fir tree) is more closely related to the stimulus at the top (e.g., a pyramid). This task requires participants to recognize all three items and compare their corresponding semantic information. The test can be conducted in either word or picture modalities to assess verbal or nonverbal semantic memory, respectively. The Kissing and Dancing Test (KDT) later extended the PPT to test actions/verbs (Bak & Hodges, 2003).

Our primary goal was to adapt the PPT and the KDT to the population of Mainland China, which is currently lacking. Certain PPT items are culturally specific, and various norms have been developed for populations other than those of the United Kingdom, including Spain (Gudayol-Ferre et al., 2008; Rami et al., 2008) and French Quebec (Callahan et al., 2010). The cultural differences between Mainland China and Western countries are so numerous that a significant proportion of the items are not suitable for Chinese speakers (e.g., windmill–tulip, nun–church). Thus, the Chinese adaptation discarded these items and added new items. Furthermore, we collected normative data on the new item sets. We considered gender, age, and level of education when obtaining the normative data because these factors can affect semantic memory performance (e.g., Brickman et al., 2000; Capitani, Laiacona, & Barbarotto, 1999; Da Silva, Peterson, Faisca, Ingvar, & Reis, 2004; Laws, 1999; but see Perlmuter & Tun, 1987). The recent Spanish and Quebec French adaptations of the PPT also showed that the level of education affected PPT performance (Callahan et al., 2010; Gudayol-Ferre et al., 2008; Rami et al., 2008).

Although the PPT and the KDT have been widely used to determine potential semantic deficits, they might not assess semantic memory comprehensively. A careful inspection of the PPT reveals that most of its items were constructed based on associative relationships (e.g., pyramid–palm tree, windmill–tulip), and in a few items targets and correct choices were from the same semantic category (i.e., the classic way of manipulating semantic similarity). One recent development in this line of research is that the semantic system is multidimensional with potential organizational principles including the modalities of knowledge and semantic categories. Neuropsychological evidence has shown that brain damage can selectively impair specific modalities of semantic knowledge such as visual color, sound, function, or manipulation knowledge, whereas other types of knowledge are relatively spared (e.g., Beauchamp, Lee, Haxby, & Martin, 2002; Boronat et al., 2005; Buxbaum & Saffran, 2002; James & Gauthier, 2003; Kellenbach, Brett, & Patterson, 2001; Miceli et al., 2001; Noppeney, Josephs, Kiebel, Friston, & Price, 2005; Noppeney & Price, 2002; Oliver & Thompson-Schill, 2003; Pulvermüller & Hauk, 2006; Tessari, Canessa, Ukmar, & Rumiati, 2007). Studies have also found selective impairments of particular semantic categories such as animals, tools, or actions (Bak & Hodges, 2003; Bi, Han, Shu, & Caramazza, 2007; Caramazza

& Shelton, 1998; Lin, Guo, Han, & Bi, 2010; Sacchett & Humphreys, 1992; Warrington & McCarthy, 1983, 1987). An additional distinction is that specific types of concepts might be organized based on different principles such as similarity versus associative relationships (Crutch, 2006; Crutch, Ridha, & Warrington, 2006; Crutch & Warrington, 2005, 2007). Neuroimaging studies have echoed these findings by showing that different modalities and categories of semantic knowledge activate different brain regions (e.g., Chao, Haxby, & Martin, 1999; for reviews, see also Gerlach, 2007; and Joseph, 2001). In light of these advances, multimeasurement approaches have been proposed to better capture the specificity of these deficits (e.g., Lambon Ralph, Moriarty, Sage, & York Speech Therapy Interest Group, 2002; Martin, Schwartz, & Kohen, 2006).

Therefore, our other aim was to develop a set of semantic tests that complement the PPT and the KDT in assessing semantic memory more comprehensively. Semantic memory organization has additional dimensions besides the associative relationships captured in the PPT. Crutch and Warrington (2005, 2007) equated the semantic similarity of concrete objects with taxonomic organization (e.g., cat–frog vs. table). Worth noting is that the distinction between taxonomy and association is not clear. While associated concepts may not share many semantic features (e.g., windmill–tulip), objects that belong to the same category are usually associatively related (for associative and similarity ratings of categorically related concrete nouns, see Zhang, Han, & Bi, 2012). In other words, many items in taxonomic tests are unavoidably associated. Nevertheless, taxonomic tests may offer information beyond the PPT: if a participant loses associative knowledge, he or she may fail the PPT test but still exhibit good performance on a taxonomic test, which suggests the preservation of some aspects of semantic memory. Related to, yet not completely overlapping with, the taxonomy organization are the similarities along various modality-specific semantic properties (i.e., features) such as form, color, sound, motion, function, and the way of manipulation. Therefore, we selected taxonomy, function, and manipulation as the semantic relationships that complement the associative relationships measured by the PPT. We chose function and manipulation properties rather than exhausting all potential modality-specific dimensions because well-documented patient and imaging findings have reported dissociations between these two modalities (e.g., Boronat et al., 2005; Buxbaum et al., 2002) and because they are essential to patients' daily functioning.

To summarize, this study seeks to (a) adapt the PPT and the KDT to the Mainland Chinese population by replacing culturally inappropriate items and collecting normative data and (b) develop three tests that complement the PPT and the KDT by assessing other types of semantic knowledge. We explored the potential relationship among these semantic tests by conducting various correlational and dissociation analyses.

METHODS

Participants

Ninety-six healthy Chinese-speaking adults were recruited from various urban communities in Shanghai, China, including 36 men (38%) and 60 women (62%).

Self-reported medical and psychiatric histories were used to ensure that all participants were free of neurological diseases, psychiatric illnesses, head injuries, and strokes. The mean age of participants was 52.9 years old (*SD* = 16.78 years, range = 20–80 years), and their mean level of education was 11.6 years (*SD* = 3.40 years, range = 3.0–17.5 years). There were no significant sex differences with regard to either of these variables (men: mean = 52.1 years old, *SD* = 17.9 years; women: mean = 53.4 years old, *SD* = 16.2 years; $t < 1$; men: mean = 12.5 years of education, *SD* = 3.2 years; women: mean = 11.1 years of education, *SD* = 3.4 years; $t = 1.82$, $p = .07$). Note that the range of age and level of education was large enough to collect normative data.

Materials and procedure

Five semantic tests were developed. All tests have the same trial structure as the picture version of the PPT in which each trial consists of three pictures printed on one sheet with one target picture on top and two alternative pictures below. These tests explicitly instruct participants to select the bottom picture that is most closely related to the top picture via a specific relationship: object association matching (PPT, Chinese adaptation [PPTc]), object taxonomy matching (OTM), object function matching (OFM), object manipulation matching (OMM), or action matching (KDT, Chinese adaptation [KDTc]). All action pictures were line drawings (including original line drawings in the KDT and self-developed ones). For object pictures, those from the original PPT were line drawings and the self-developed ones were black-and-white photographs of common objects. We used photographs for the new items because (a) line drawings of these items were not as easily acquired as photographs, (b) novel line drawings may differ from the original PPT items because styles vary across artists, and (c) photographs may be more recognizable than line drawings and better reduce the variance that originates from the picture-perception stage. We also collected the word frequency data (Sun, Huang, Sun, Li, & Xing, 1997) that corresponded to the picture stimuli within each test (Table 1). This database was constructed based on each word's frequency of occurrence per one million words.

PPTc. The PPTc contains the original 52 items from the PPT (Howard & Patterson, 1992) and 28 additional items suitable to Chinese culture constructed using the same association principle as the other items. This was done to ensure that there would be enough good items to select from in case too many original items were inappropriate. Participants decide which of the bottom pictures (e.g., a palm tree and a pine tree) is more closely related to the top picture (e.g., a pyramid). The instructions were as follows: "Three pictures are oriented in a triangle on each sheet. Please choose which of the two bottom pictures is more related to the top picture."

KDTc. The KDTc has the same structure as the PPTc, and action pictures are used. The KDTc includes the original 52 items from the KDT (Bak & Hodges, 2003) and 28 additional items constructed for the Chinese culture. As in the PPTc, participants select which of the two bottom actions (e.g., blowing and drinking) is

Table 1. *Mean accuracy, standard deviation, and mean log word frequency of each test*

	PPTc	KDTc	OTM	OFM	OMM
No. of participants	96	96	79	79	82
Mean accuracy	94%	92%	92%	93%	86%
SD	0.06	0.08	0.06	0.07	0.10
Range	0.66–1	0.54–1	0.66–1	0.67–1	0.59–1
Mean log word frequency	0.90	1.06	0.72	0.64	0.54
SD	0.70	0.85	0.53	0.53	0.46

Note: PPTc, Pyramids and Palm Trees Test, Chinese adaptation; KDTc, Kissing and Dancing Test, Chinese adaptation; OTM, object taxonomy matching test; OFM, object function matching test; OMM, object manipulation matching test.

more closely related to the top action (e.g., eating). The instructions were identical to those of the PPTc.

OTM test. This test is composed of 38 items, and participants judge which of the two bottom pictures (e.g., a pear and a bean) is more similar to the top picture (e.g., an apple). These pictures were constructed such that target and response were always from the same semantic (coordinate or subordinate) category. The instructions were as follows: “Three pictures are oriented in a triangle on each sheet. Please choose which of the two bottom pictures is more similar to the top picture.”

OFM test. This test includes 36 trials. Participants choose which of the bottom pictures (e.g., a drill and a video camera) has a function that is more similar to the top picture (e.g., a camera). The instructions were as follows: “Three pictures are oriented in a triangle on each sheet. Please choose which of the two bottom pictures has a function that is more similar to that of the top picture. For example, chopsticks and spoons are both used to eat.”

OMM test. This test has 22 trials. Participants choose which of the bottom pictures (e.g., a stapler and an iron) is manipulated in a way that is more similar to that of the top picture (e.g., a saw). The instructions were as follows: “Three pictures are oriented in a triangle on each sheet. Please choose which of the two bottom pictures is more similar to the top picture in terms of how they are manipulated. For example, a broom is held with both hands and moved back and forth.”

To prevent participants from using response strategies that rely on low-level visual cues, we designed some trials to have visually similar but incorrect choices (e.g., “clip–clothespin/pencil,” where pencil is the correct choice in the OTM test). Participants had three practice trials in the PPTc, the first test of the testing session, and two practice trials in the OMM, a relatively more difficult test among the cohorts. There was no practice trial in the KDTc, the OTM, or the OFM.

Due to various pragmatic reasons, not all participants completed all five tests. The PPTc and the KDTc were administered to all 96 participants. The OTM, the OFM, and the OMM were administered to 79, 79, and 82 participants, respectively. All participants completed the tests in the following order: PPTc, KDTc, OTM, OMM, and OFM. Participants required approximately 25–35 min to finish all the tests. Following the conventions of clinical neuropsychological assessments, we did not impose a time limit. Furthermore, we did not provide participants with feedback during the testing session; if participants stated that they did not know the answer, we advised them to guess. We repeated the instructions to participants who had difficulty remembering them when requested throughout the task.

RESULTS

We conducted analyses to construct the normative data set and then explored the relationships among the different tests.

Normative data development

Items that yielded error rates over 20% on the PPTc and the KDTc (i.e., those that were responded to correctly by less than 80% of participants) were considered as inappropriate for Chinese participants and were eliminated from further analyses. We adopted a more liberal inclusion criterion than the original PPT and KDT tests to preserve as many items as possible so that the adapted tests were more comparable to the originals. We checked the items with agreement rates below 80% and observed that most of these items were not familiar to Chinese people (e.g., acorn–donkey/pig). In addition, we examined the frequency data of the PPT and the KDT items that were discarded, those that were retained, and the new items. The results showed that for the PPTc, there were significant differences among the word frequencies of the retained PPT items, the new items, and the discarded items, F (2, 237) = 7.6, $p < .001$, such that the frequencies of the retained PPT items were higher than those of the others (post hoc comparison ps $< .01$); for the KDTc, the differences among the word frequencies of the retained KDT items, the new items, and the discarded items were also significant, F (2, 237) = 6.5, $p < .01$, such that the frequencies of the retained KDT items and the new ones were higher than those of the discarded items (post hoc comparison ps $< .01$). Therefore, these results confirmed our observation that the discarded items were not suitable for Chinese participants.

We did not remove any items from the OTM, OFM, and OMM tests because they were developed only using items familiar to Chinese participants. The “culture adaptation” only applies to the PPTc and the KDTc, which were originally developed for British participants. Items of the OTM and OFM tests were rarely below the 80% criterion. The overall performance on the OMM test was lower (e.g., certain items yielded error rates above 20%) because the typical way of manipulating an object is relatively more ambiguous than other object properties such as object function. Thus, caution is needed when interpreting the participants’ performance on the OMM test by (a) referencing the norm acquired below and (b) considering other tests that involve manipulation knowledge such as object use.

Table 2. *Results of the stepwise regression analyses*

	PPTc		OTM		OMM	
	t	*p*	*t*	*p*	*t*	*p*
Edu	1.70	.09	2.09*	.04	2.48*	.02
Age	0.19	.85	0.23	.82	−2.02*	.05
Edu × Age	−3.23**	.00	0.31	.76	−0.35	.73
Gender	−0.32	.75	0.48	.64	−0.83	.41

Note: PPTc, Pyramids and Palm Trees Test, Chinese adaptation; OTM, object taxonomy matching test; OMM, object manipulation matching test; Edu, level of education.
$*p < .05$. $**p < .01$.

Thus, 64 PPTc items and 67 KDTc items as well as all the items of the OTM, OFM, and OMM tests were formally analyzed and are presented below. Each participant's score for each test was the percentage of correct responses. Table 1 lists the mean accuracy and variance for each test.

To appropriately categorize the normative data along sociodemographic indices, we conducted a regression analysis to identify the sociodemographic variables that influence semantic performance. Stepwise regression analyses were performed using age, sex, and level of education as predictors. Age and level of education were entered as continuous variables in the analysis. Women and men were coded as 0 and 1, respectively. We also included the age by education interaction as a predictor because a previous study reported that these two variables significantly interacted (Gudayol-Ferre et al., 2008). The interaction was centered such that each value was replaced with its difference from the mean to reduce multicollinearity. We also transformed the test scores, 2 arcsin (score$^{1/2}$), for the regression analyses because these distributions were skewed.

The regression models for the KDTc and OFM tests were not significant; Table 2 shows the regression results for the PPTc as well as the OTM and OMM tests. The education by age interaction significantly predicted PPTc performance ($t = –3.23$, $p < .01$), and the level of education effect was marginally significant ($t = 1.70$, $p = .09$). Therefore, we conducted an additional independent-samples *t* test to compare the performance across different age (20–55 vs. 56+ years) and level of education groups (3–12 vs. 13+ years). We found that participants less than 56 years old with high levels of education performed significantly better than those with less education ($t = –3.72$, $p < .001$); of those with less education, older participants scored significantly higher than younger participants ($t = –2.23$, $p < .05$). Level of education significantly predicted OTM test performance ($t = 2.09$, $p < .05$). Furthermore, both level of education ($t = 2.48$, $p < .05$) and age ($t = –2.02$, $p < .05$) significantly influenced the OMM test score. Additional *t* tests confirmed that there were significant differences between the two education groups with regard to the PPTc ($t = –2.13$, $p < .05$) and the OMM test ($t = –3.67$, $p < .001$) performance. There were marginally significant differences in the OMM scores between the two age groups ($t = 1.88$, $p = .06$).

Gender did not predict the performance on any test. This lack of a result might be because men and women were unequally represented in the current research. The gender distribution by age and level of education was the following: 45% men in group of age 20–55, years of education 3–12; 42% men in age 20–55, years of education 13+; 24% men in age 56+, years of education 3–12; and 50% men in age 56+, years of education 13+ (Table 3). Therefore, we still considered gender to be a factor along which the norm was stratified.

In light of the regression and *t* test results, we stratified the normative data by gender, age (i.e., 20–55 years vs. 56+ years), and level of education (i.e., 3–12 years vs. 13+ years). Table 3 lists the resulting norms. The scores were grouped according to gender, age, and level of education, and cutoff scores were set at 5% (Z score = –1.65) to determine pathological performance.

Correlational and dissociation analyses

To explore the response patterns across different semantic tests, we conducted both correlational analyses and dissociation statistics (Crawford & Garthwaite, 2005). We were interested in whether at the group level the performance correlated across various semantic tasks and whether at the case level the performance dissociated. These two analyses converged to explore specific shared and independent cognitive components.

Table 4 shows the correlation matrix derived using the Pearson *r*. All tasks were significantly correlated with each other.

The dissociation statistical program we used (Crawford & Garthwaite, 2005) tests whether an individual's score on one task significantly differs from his/her score on another after considering the group performance on these two tasks. We examined the contrasts between all tasks to see whether any participant showed a dissociated performance (i.e., performed significantly better on one task compared with another; $p < .05$, two tailed). These contrasts showed that 22 cases significantly dissociated: one case in the PPTc versus the KDTc, the PPTc versus the OMM test, the KDTc versus the OFM test, and the KDTc versus the OMM test; two in the KDTc versus the OTM test and the OFM versus the OMM test; three in the PPTc versus the OTM test, the PPTc versus the OFM test, and the OTM versus the OMM test; and five in the OTM versus the OFM test. Given that we conducted 813 comparisons, we expected 5% (41 cases) of the contrasts to be significant ($p < .05$, two tailed). That is, the 22 dissociations we observed were expected due to chance alone.

DISCUSSION

We developed Chinese adaptations of the PPT and the KDT as well as three additional semantic tests that were constructed along different types of semantic knowledge and collected normative data in a group of healthy, Chinese-speaking participants. We found that level of education and age affected semantic performance in complex ways. Furthermore, the performance on different semantic tests

Table 3. *Test norms stratified by gender, age (20–55 years vs. 56+ years), and level of education (3–12 years vs. 13+ years)*

	3–12 Years of Education				13+ Years of Education			
	Age 20–55 Years		Age 56+ Years		Age 20–55 Years		Age 56+ Years	
	F ($N = 11$)	M ($N = 9$)	F ($N = 28$)	M ($N = 9$)	F ($N = 11$)	M ($N = 8$)	F ($N = 10$)	M ($N = 10$)
PPTc								
Mean	0.90	0.89	0.94	0.92	0.97	0.95	0.94	0.94
SD	0.06	0.11	0.05	0.07	0.04	0.05	0.04	0.04
5% cutoff	0.80	0.71	0.85	0.80	0.91	0.87	0.87	0.87
KDTc								
Mean	0.91	0.85	0.92	0.92	0.92	0.95	0.91	0.91
SD	0.09	0.16	0.06	0.04	0.07	0.04	0.05	0.07
5% cutoff	0.76	0.58	0.82	0.85	0.81	0.88	0.83	0.79
OTM								
Mean	0.91	0.90	0.91	0.91	0.91	0.95	0.92	0.93
SD	0.05	0.08	0.05	0.04	0.10	0.04	0.06	0.05
5% cutoff	0.82	0.76	0.84	0.85	0.74	0.89	0.83	0.86
OFM								
Mean	0.90	0.90	0.94	0.91	0.97	0.98	0.91	0.94
SD	0.10	0.07	0.05	0.08	0.04	0.02	0.07	0.06
5% cutoff	0.72	0.78	0.87	0.77	0.91	0.96	0.79	0.85
OMM								
Mean	0.86	0.82	0.83	0.83	0.97	0.93	0.85	0.89
SD	0.12	0.07	0.11	0.09	0.05	0.07	0.04	0.11
5% cutoff	0.66	0.71	0.64	0.69	0.89	0.82	0.79	0.71

Note: The 5% cutoff score corresponds to a Z score of –1.65. F, female; M, male; PPTc, Pyramids and Palm Trees Test, Chinese adaptation; KDTc, Kissing and Dancing Test, Chinese adaptation; OTM, object taxonomy matching test; OFM, object function matching test; OMM, object manipulation matching test.

Table 4. *Pearson's correlation coefficients among tests*

	PPTc	KDTc	OTM	OFM	OMM
PPTc					
KDTc	.76**				
OTM	.42**	.34**			
OFM	.40**	.35**	.38**		
OMM	.33**	.28*	.29**	.36**	

Note: PPTc, Pyramids and Palm Trees Test, Chinese adaptation; KDTc, Kissing and Dancing Test, Chinese adaptation; OTM, object taxonomy matching test; OFM, object function matching test; OMM, object manipulation matching test.
$^{**}p < .01$.

was largely correlated with each other. We discuss these aspects of the results below.

Normative data of the semantic tests

The primary outcomes of our study are the Chinese adaptation of the PPT and the KDT. As discussed in the Introduction, the large cultural differences between China and the West render direct translations of these popular tests as highly inappropriate. We not only excluded culturally inappropriate items (e.g., windmill–tulip; nun–church) but also added items that were specifically suited to the Chinese population (e.g., chopstick–bowl).

Furthermore, the normative data collected for all semantic tests were stratified by gender, age (20–55 vs. 56+ years), and level of education (3–12 vs. 13+ years) based on the results of the regression analyses. Following Callahan et al. (2010), we set the individual cutoff score for each demographic cluster at 5%, which can be used to identify pathological deficits. Furthermore, this method takes into account participant gender, age, and level of education and is therefore more accurate than the arbitrary 90% cutoff that Howard and Patterson (1992) originally proposed. It is important that, because our participants were largely from an urban area, future researchers should be cautious when applying the tests to rural Chinese populations.

The effects of sociodemographic variables on semantic tests

Regression analyses revealed that level of education significantly influenced performance on the OTM and OMM tests, and age significantly predicted the OMM test score. These variables modulated each other in predicting PPTc performance. We did not observe any gender effects.

The positive effects of education on various semantic tasks are consistent with previous studies (e.g., Callahan et al., 2010; Da Silva et al., 2004; Rami et al., 2008). The age effect was significant only with regard to the OMM test, most likely because this test was the most difficult (as defined by the lowest mean response accuracy), while this effect was too weak to emerge in the "easier" tests (e.g., the PPTc) due to a ceiling effect (see also Callahan et al., 2010; Rami et al., 2008). Consistent with the results of the Spanish study (Gudayol-Ferre et al., 2008), we observed an interaction effect of age by education on the PPTc. Younger participants (i.e., 20–55 years old) with more education showed better performance than their counterparts. A counterintuitive result was that younger participants with less education (i.e., 3–12 years) scored lower on the PPTc than elder participants with less education. This result might be because semantic associations can be constructed through life experience when education is lacking.

There was no gender effect in our study. Although this result is consistent with some PPT studies (e.g., Callahan et al., 2010; Gudayol-Ferre et al., 2008; Rami et al., 2008), others have reported gender effects with regard to the processing of certain semantic categories. Healthy females have performed better than males with regard to naming living things, whereas males have performed better when naming nonliving things (Capitani et al., 1999; Laws, 1999). Familiarity has been used to explain this gender effect (Albanese, Capitani, Barbarotto, & Laiacona, 2000). Although the current study did not manipulate the semantic category variable, there were opportunities to observe potential gender differences because the OMM and OFM test items that were exclusively nonliving artifacts. Therefore, gender effects might only be observable in naming, but not comprehensive, tasks. Another possible explanation for the lack of a gender effect was that men and women were not equally represented in the current research. Therefore, gender should be considered when referring to the normative data.

Comparisons across different semantic tests

Two lines of results emerged from the correlational and dissociation analyses with regard to performance across the different semantic tests. The correlations among all tests were significant, and the highest degree of correlation was observed between the PPTc and the KDTc. These correlations indicate that the semantic tests share components. As many researchers have proposed (e.g., Patterson et al., 2007), one possibility is that various semantic attributes are represented distributedly and are bound together in specific brain regions. Different levels of semantic-task performance (e.g., the PPT) among healthy people might primarily originate from variations in the functioning of these bound regions. Although the semantic tests were designed to assess different types of semantic knowledge, other semantic aspects might be automatically activated and thereby affect performance. For instance, functional knowledge (e.g., keys are for unlocking) might assist manipulation judgments (e.g., turning) for particular objects. Nevertheless, the observed across-task dissociations did not occur more often than was expected by chance. In addition to the explanations provided above, accuracy might not be a sensitive enough measure in healthy participants, especially those who are young; therefore, we might not have detected the potential dissociations among the different semantic tests. Finally, the modulation patterns of various types of individual experiences on semantic organization might be much more complicated than originally assumed. For instance, the manipulation of a tool might be more related to one's life experience, whereas education more strongly influences the categorical knowledge of tools. Future studies should more carefully consider different types of life-experience measures with regard to semantic organization. Nevertheless, given the theoretical motivations and the empirical evidence reported in the literature concerning the dissociations among different aspects of semantic knowledge (e.g., Buxbaum et al., 2002; Laiacona & Caramazza, 2004), the tests that we developed may prove useful in future clinical settings to detect semantic impairments specific to one or more aspects of properties.

CONCLUSION

We adapted the PPT and the KDT for the population of Mainland China by adopting culturally appropriate items and collecting normative data. Furthermore, we developed three additional semantic tests to examine different aspects of semantic knowledge. We observed highly significant correlations across these semantic tests, which suggest that the completion of these tasks involves common components. Whether the developed tests allow clinicians to detect semantic impairments that complement the PPT and the KDT warrants future examination.

APPENDIX A

Table A.1, Table A.2, Table A.3, Table A.4, and Table A.5 present the stimuli developed by the present study. The words in italic represent the correct choices. This list can also be accessed via our website (http://psychbrain.bnu.edu.cn/home/yanchaobi/supplement/Appendix.docx).

Table A.1. *Pyramids and Palm Trees Test, Chinese adaptation*

	Target	Choice 1	Choice 2
	New Items		
1	Frog	*Tadpole*	Lion
2	Envelope	*Stamp*	Camera
3	Belt	*Pants*	Watch
4	Pavilion	*Bench*	Building
5	Barrel	*Well*	Monument
6	Railway	*Train*	Bus
7	Parachute	Ship	*Airplane*
8	Washbasin	Flashlight	*Tap*
9	Key	Steelyard	*Lock*
10	Tie	*Suit*	Vest
11	Blackboard	*Desk*	Tent

Table A.1 (*cont.*)

	Target	Choice 1	Choice 2
		New Items	
12	Stop sign	*Bus*	Airplane
13	Album	Lighter	*Camera*
14	Stapler	*Desk*	Sound box
15	Monkey	*Peach*	Carrot
16	Film	*Video camera*	Microwave oven
17	Elevator	*Building*	Tower
18	Bowl	*Chopstick*	Clip
19	Umbrella	Pants	*Rain boots*
20	Quilt	Mask	*Pillow*
21	Wedding dress	*Ring*	Bracelet
22	Bathtub	Hairdryer	*Tap*
23	Button	*Pin*	Bench
24	Book shelf	*Desk*	Parallel bars
25	Straw	Disc	*Cup*
26	Mouse	Abacus	*Computer*
27	Bullet	*Pistol*	Longbow
28	TV	Flashlight	*Remote control*
		Discarded Items	
1	Nun	*Church*	House
2	Milk	*Cow*	Bull
3	Mask	*Clown*	Mayor
4	Thimble	*Needle*	Bobbin
5	Ticket	Car	*Bus*
6	Eskimo	*Dome*	Building
7	Pyramid	*Palm tree*	Pine tree
8	Windmill	Daffodils	*Tulips*
9	Carrot	Lamb	*Donkey*
10	Ring	*Middle finger*	Thumb
11	Soldier	Church	*Castle*
12	Rocket	Pentacle	*Sky*
13	Pin	Girl	*Baby*
14	Acorn	Donkey	*Pig*
15	Sketchpad	Table	*Desk*
16	Eskimo	Boat	*Kayak*

Note: Only the new items and the discarded items of the original Pyramids and Palm Trees Test are listed.

Table A.2. *Kissing and Dancing Test, Chinese adaptation stimuli*

	Target	Choice 1	Choice 2
	New Items		
1	Mirroring	Drinking	*Combing*
2	Teaching	*Learning*	Eating
3	Beating	*Kicking*	Pulling
4	Buckling	Waving	*Zipping*
5	Cooking	*Frying*	Dropping
6	Sweeping the floor	*Pouring trash*	Playing piano
7	Undressing	*Bathing*	Falling
8	Breaking	*Tearing*	Packing
9	Throwing a snowball	*Flapping bowling*	Boxing
10	Camping	*Barbecuing*	Drying clothes
11	Brushing teeth	*Face Washing*	Wearing shoes
12	Raking	Driving	*Shoveling*
13	Running	Singing	*Walking*
14	Frying	*Eating*	Playing baseball
15	Saluting	*Stepping*	Kneeling
16	Writing	Salvaging	*Erasing*
17	Sweating	*Running*	Raising a hand
18	Playing piano	*Dancing*	Boxing
19	Pushing	*Pulling*	Plugging
20	Abrading	Watering	*Cutting*
21	Bathing	*Sleeping*	Saluting
22	Planting	Sculpting	*Picking flowers*
23	Standing	*Kneeling*	Eating
24	Climbing	Crying	*Slipping*
25	Bathing	*Drying*	Peeling
26	Singing	Climbing	*Dancing*
	Discarded Items		
1	Yawning	Jumping	*Sleeping*
2	Watering	*Pouring*	Peeling
3	Fighting	*Shooting*	Riding
4	Washing	*Ironing*	Knotting
5	Skiing	*Skating*	Swimming
6	Shutting	*Opening*	Giving
7	Buying	*Robbing*	Teaching
8	Painting	*Drawing*	Touching
9	Knocking	*Greeting*	Cutting
10	Kissing	*Dancing*	Running
11	Falling	*Slipping*	Bouncing

Note: Only the new items and the discarded items of the original Kissing and Dancing Test are listed.

Table A.3. *Object taxonomy matching test stimuli*

	Target	Choice 1	Choice 2
1	Rooster	Eagle	*Duck*
2	Corn	Garlic	*Wheat*
3	Banana	Onion	*Orange*
4	Owl	*Bat*	Zebra
5	Chili	*Ginger*	Grape
6	Apple	*Pear*	Green bean
7	Book shelf	Parallel bar	*Desk*
8	Potato	*Sweet potato*	Apple
9	Saw	*Wrench*	Peg
10	Spider	Shrimp	*Scorpion*
11	Penguin	*Polar bear*	Elephant
12	Dolphin	*Whale*	Cockroach
13	Watermelon	*Apple*	Walnut
14	Carrot	*Lotus root*	Eggplant
15	Pan	Hair dryer	*Bailer*
16	Blackboard	*Book shelf*	Tent
17	Abacus	Needle	*Steelyard*
18	Bathtub	*Sink*	Book shelf
19	Peanut	*Walnut*	Potato
20	Ginger	Cauliflower	*Garlic*
21	Pumpkin	Pear	*Cabbage*
22	Camel	Squirrel	*Elephant*
23	Antelope	Elephant	*Deer*
24	Ruler	*Steelyard*	Spoon
25	Pencil box	Anchor	*Envelope*
26	Building	Aircraft	*Castle*
27	Film	*Disc*	Mask
28	Crab	Dragonfly	*Shrimp*
29	Cat	*Tiger*	Frog
30	Well	*Fence*	Tap
31	Ship	*Car*	Lift
32	Flashlight	*Candle*	Glove
33	Rat	*Squirrel*	Deer
34	Pavilion	Chimney	*Tower*
35	Camera	*Video camera*	Microwave oven
36	Clip	Clamp	*Pencil*
37	Watch	*Alarm clock*	Cup
38	Gloves	Pillow	*Pant*

Table A.4. *Object function matching test stimuli*

	Target	Choice 1	Choice 2
1	Brush	Abacus	*Roller brush*
2	Bookcase	*Cabinet*	Stop sign
3	Album	Thermometer	*Tape*
4	Kettle	Remote control	*Pot*
5	Clip	*Stapler*	Bailer
6	Monument	Bathtub	*Statue*
7	Bracelet	*Necklace*	Envelope
8	Chimney	*Range hood*	Bookshelf
9	Speaker	*Sound box*	Sliding board
10	Blackboard	Tower	*Stop sign*
11	Camera	Drill	*Video camera*
12	Tent	*Building*	Statue
13	Glasses	*Magnifying glasses*	Flashlight
14	Electric fan	*Air conditioning*	Microwave oven
15	Castle	Building	*Tower*
16	Fishing rod	Zip	*Fishing net*
17	Broom	*Mop*	Bailer
18	Chopstick	*Fork*	Pliers
19	Train	*Bus*	Wall
20	Abacus	Brush	*Calculator*
21	Match	Fork	*Lighter*
22	Umbrella	*Raincoat*	Belt
23	Button	Valve	*Zip*
24	Fence	*Wall*	Bus
25	Knife	*Scissors*	Whip
26	Candle	Pot	*Flashlight*
27	Bucket	Pavilion	*Bathtub*
28	Heating	*Air conditioning*	Stop sign
29	Watch	Bag	*Alarm clock*
30	Pan	Watch	*Microwave oven*
31	Disc	*Tape*	Book
32	Elevator	*Stair*	Wall
33	Fan	*Electrical fan*	Iron
34	Scarf	Watch	*Hat*
35	Spoon	Eraser	*Straw*
36	Ax	Wrench	*Saw*

Table A.5. *Object manipulation matching test*

	Target	Choice 1	Choice 2
1	Scarf	*Necklace*	Camera
2	Vest	*Raincoat*	Pillow
3	Pan	*Cup*	Hat
4	Spoon	*Bailer*	Pot
5	Saw	Stapler	*Iron*
6	Bracelet	*Watch*	Bowl
7	Brush	Mouse	*Eraser*
8	Remote control	Glasses box	*Flashlight*
9	Belt	*Watch*	Necklace
10	Envelope	*Wallet*	Eraser
11	Straw	Ax	*Pipe*
12	Hammer	*Ax*	Wrench
13	Hairdryer	Saw	*Pistol*
14	Clamp	*Forceps*	Fork
15	Kettle	Pipe	*Drill*
16	Schoolbag	Pillow	*Vest*
17	Scissor	*Pliers*	Knife
18	Wrench	Spoon	*Valve*
19	Album	Envelope	*Calendar*
20	Alarm clock	Video camera	*Microwave oven*
21	Fishing rod	*Whip*	Pencil
22	Broom	*Rake*	Fork

ACKNOWLEDGMENTS

The first two authors contributed equally to this work and should both be considered as first authors. This work was supported by 973 Program (2013CB837300), Major Project of National Social Science Foundation (11&ZD186), NSFC (31171073, 31222024, and 31271115), and the Science and Technology Commission of Shanghai Municipality (Grants 08411951000 and 09DZ1950400). We thank the anonymous reviewers for their helpful comments on an earlier version of this manuscript, and we are grateful to all of the participants in the study.

REFERENCES

Albanese, E., Capitani, E., Barbarotto, R., & Laiacona, M. (2000). Semantic category dissociations, familiarity and gender. *Cortex, 36*, 733–746.

Bak, T. H., & Hodges, J. R. (2003). Kissing and dancing—A test to distinguish the lexical and conceptual contributions to noun/verb and action/object dissociation: Preliminary results in patients with frontotemporal dementia. *Journal of Neurolinguistics, 16*, 169–181.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron, 34*, 149–159.

Bi, Y., Han, Z., Shu, H., & Caramazza, A. (2007). Nouns, verbs, objects, actions, and the animate/inanimate effect. *Cognitive Neuropsychology, 24*, 485–504.

Boronat, C. B., Buxbaum, L. J., Coslett, H. B., Tang, K., Saffran, E. M., Kimberg, D. Y., et al. (2005). Distinctions between manipulation and function knowledge of objects: Evidence from functional magnetic resonance imaging. *Cognitive Brain Research, 23*, 361–373.

Brickman, A. M., Paul, R. H., Cohen, R. A., Williams, L. M., MacGregor, K. L., Jefferson, A. L., et al. (2000). Category and letter verbal fluency across the adult lifespan: Relationship to EEG theta power. *Archives of Clinical Neuropsychology, 20*, 561–573.

Buxbaum, L. J., & Saffran, E. M. (2002). Knowledge of object manipulation and object function: Dissociations in apraxic and nonapraxic subjects. *Brain and Language, 82*, 179–199.

Buxbaum, L. J., Schwartz, M. F., & Carew, T. G. (1997). The role of semantic memory in object use. *Cognitive Neuropsychology, 14*, 219–254.

Callahan, B. L., Macoir, J., Hudon, C., Bier, N., Chouinard, N., Cossette-Harvey, M., et al. (2010). Normative data for the Pyramids and Palm Trees Test in the Quebec–French population. *Archives of Clinical Neuropsychology, 25*, 212–217.

Capitani, E., Laiacona, M., & Barbarotto, R. (1999). Gender affects word retrieval of certain categories in semantic fluency tasks. *Cortex, 35*, 273–278.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience, 10*, 1–34.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*, 913–919.

Crawford, J. R., & Garthwaite, P. H. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: Evaluation of alternatives using Monte Carlo simulations and revised tests for dissociations. *Neuropsychology, 19*, 318–331.

Crutch, S. J. (2006). Qualitatively different semantic representations for abstract and concrete words: Further evidence from the semantic reading errors of deep dyslexic patients. *Neurocase, 12*, 91–97.

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. (2007). Semantic priming in deep-phonological dyslexia: Contrasting effects of association and similarity upon abstract and concrete word reading. *Cognitive Neuropsychology, 24*, 1–20.

Da Silva, C. G., Petersson, K. M., Faisca, L., Ingvar, M., & Reis, A. (2004). The effects of literacy and education on the quantitative and qualitative aspects of semantic verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 26*, 266–277.

Gerlach, C. (2007). A review of functional imaging studies on category specificity. *Journal of Cognitive Neuroscience, 19*, 296–314.

Gudayol-Ferre, E., Lara, J. P., Herrera-Guzman, I., Bohm, P., Rodes, E., Ansaldo, A. I., et al. (2008). Semantic memory as assessed by the Pyramids and Palm Trees Test: The impact of sociodemographic factors in a Spanish-speaking population. *Journal of the International Neuropsychological Society, 14*, 148–151.

Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test: A test for semantic access from words and pictures*. Bury St. Edmunds: Thames Valley Test Company.

James, T. W., & Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology, 13*, 1792–1796.

Joseph, J. E. (2001). Functional neuroimaging studies of category specificity in object recognition: A critical review and meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience, 1*, 119–136.

Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive, Affective, & Behavioral Neuroscience, 1*, 207–221.

Laiacona, M., & Caramazza, A. (2004). The noun/verb dissociation in language production: Varieties of causes. *Cognitive Neuropsychology, 21*, 103–123.

Lambon Ralph, M. A. L., Moriarty, L., Sage, K., & York Speech Therapy Interest Group. (2002). Anomia is simply a reflection of semantic and phonological impairments: Evidence from a case-series study. *Aphasiology, 16*, 56–82.

Laws, K. R. (1999). Gender affects naming latencies for living and nonliving things: Implications for familiarity. *Cortex, 35*, 729–733.

Lin, N., Guo, Q., Han, Z., & Bi, Y. (2010). Motor knowledge is one dimension for concept organization: Further evidence from a Chinese semantic dementia case. *Brain and Language, 119*, 110–118.

Martin, N., Schwartz, M. F., & Kohen, F. P. (2006). Assessment of the ability to process semantic and phonological aspects of words in aphasia: A multi-measurement approach. *Aphasiology, 20*, 154–166.

Miceli, G., Fouch, E., Capasso, R., Shelton, J. R., Tomaiuolo, F., & Caramazza, A. (2001). The dissociation of color from form and function knowledge. *Nature Neuroscience, 4*, 662–667.

Noppeney, U., Josephs, O., Kiebel, S., Friston, K. J., & Price, C. J. (2005). Action selectivity in parietal and temporal cortex. *Cognitive Brain Research, 25*, 641–649.

Noppeney, U., & Price, C. J. (2002). Retrieval of visual, auditory, and abstract semantics. *NeuroImage, 15*, 917–926.

Oliver, R. T., & Thompson-Schill, S. L. (2003). Dorsal stream activation during retrieval of object size and shape. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 309–322.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*, 976–987.

Perlmuter, L. C., & Tun, P. (1987). Age and diabetes related changes in verbal fluency. *Experimental Aging Research, 13*, 9–14.

Pulvermüller, F., & Hauk, O. (2006). Category-specific conceptual processing of color and form in left fronto-temporal cortex. *Cerebral Cortex, 16*, 1193–1201.

Rami, L., Serradell, M., Bosch, B., Caprile, C., Sekler, A., Villar, A., et al. (2008). Normative data for the Boston Naming Test and the Pyramids and Palm Trees Test in the elderly Spanish population. *Journal of Clinical and Experimental Neuropsychology, 30*, 1–6.

Sacchett, C., & Humphreys, G. W. (1992). Calling a squirrel a squirrel but a canoe a wigwam: A category-specific deficit for artifactual objects and body parts. *Cognitive Neuropsychology, 9*, 73–86.

Sun, H. L., Huang, J. P., Sun, D. J., Li, D. J., & Xing, H. B. (1997). Introduction to language corpus system of modern Chinese study. In M. Y. Hu (Ed.), *Paper collection for the Fifth World Chinese Teaching Symposium* (pp. 459–466). Beijing: Peking University Press.

Tessari, A., Canessa, N., Ukmar, M., & Rumiati, R. (2007). Neuropsychological evidence for a strategic control of multiple routes in imitation. *Brain, 130*, 1111–1126.

Tulving, E. (1972). Episodic and semantic memory. In E. T. W. Donaldson (Ed.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27*, 635–657.

Warrington, E. K., & McCarthy, R. A. (1983). Category specific access dysphasia. *Brain, 106*, 859–878.

Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain, 110*, 1273–1296.

Zhang, X., Han, Z., & Bi, Y. (2012). Are abstract and concrete concepts organized differently? Evidence from the blocked translation paradigm. *Applied Psycholinguistics*. Advance online publication. doi:10.1017/S0142716412000124