The Effect of Feature Integration and Processing Depth on the Association Between Moral Concepts and Container Metaphors

Authors: Wang Congxing, Yang Yuqin, Xiong Meng, Ye Yiduo

Date: 2020-10-11T00:00:00+00:00

Abstract

The metaphor extraction hypothesis proposes a separation between the formation and extraction of metaphorical associations. Three experiments investigated the metaphorical association between moral concepts and container space and its modulation by perceptual processing depth and degree of feature integration. Experiment 1 employed a spatial Stroop paradigm, Experiments 2a and 2b employed a priming paradigm, and Experiments 3a and 3b employed a Stroop paradigm with additional task requirements. The results revealed: (1) No metaphorical association between moral concepts and container space was observed in the classic Stroop paradigm; (2) In the priming paradigm, a weak metaphorical association emerged under deeper perceptual processing; (3) In the Stroop task with higher feature integration, a strong metaphorical association was present. These findings indicate that moral concepts and container space exhibit a metaphorical association wherein moral is inside and immoral is outside, that this association demonstrates bidirectional mapping, and that it is influenced by both feature integration degree and perceptual processing depth, thereby providing support for the metaphor extraction hypothesis.

Full Text

The Influence of Feature Integration and Processing Depth on Metaphorical Association between Moral Concepts and Container Space

WANG Congxing[1]; **YANG Yuqing**[1]; **XIONG Meng**[2,3]; **YE Yiduo**[1]
([1]School of Psychology, Fujian Normal University, Fuzhou 350000, China)
([2]Department of Psychology, School of Education and Sport Sciences, Yangtze University, Jingzhou 434023, China)

([3]Department of Psychology, The University of Edinburgh, Edinburgh EH89JZ, UK)

Abstract

This study proposes the Metaphor Retrieval Hypothesis to distinguish between the formation and retrieval of metaphorical associations, investigating the metaphorical connection between moral concepts and container space through three experiments. Experiment 1 employed a spatial Stroop paradigm, Experiments 2a and 2b used a priming paradigm, and Experiments 3a and 3b utilized a Stroop paradigm with additional task demands. The results revealed: (1) No metaphorical association between moral concepts and container space was found in the classic Stroop paradigm; (2) In the priming paradigm, a weak metaphorical association emerged under deeper perceptual processing depth; (3) In the Stroop task with higher feature integration, a stronger metaphorical association was observed. These findings indicate that moral concepts and container space exhibit a metaphorical association where "moral is inside" and "immoral is outside," characterized by bidirectional mapping that is influenced by both feature integration degree and perceptual processing depth. The results provide empirical support for the Metaphor Retrieval Hypothesis.

Keywords: moral; container; metaphor; mapping; depth of processing; feature integration; metaphor retrieval

Classification Code: B842

How humans understand abstract concepts has long been a central topic in cognitive linguistics (Borghi et al., 2017; He & Chen, 2020; Wang & Yao, 2012). With the rise of embodied cognition, the understanding of abstract concept representation has undergone a significant epistemological shift. Abstract concepts can only be acquired and understood through concrete concepts and sensorimotor experiences, with metaphor serving as the intrinsic psychological mechanism underlying this constructive process (Barsalou, 2008; Gibbs, 2006; Holyoak & Stamenković, 2018; Wang & Lu, 2003; Wilson, 2002; Yin & Ye, 2014; Yin, Su, & Ye, 2013).

1.1.1 Psychological Reality and Embodied Origins

Received: 2020-05-21
Corresponding author: YE Yiduo, E-mail: yeyiduo@163.com

Moral concepts, as typical abstract concepts, have received considerable research attention. The embodied cognition perspective posits that the acquisition of moral concepts is grounded in bodily experience and constrained by physiological and environmental factors (Chen et al., 2014). Sensorimotor experiences are not isolated islands floating outside higher-level information processing; rather, moral concept formation is not merely a symbolic issue but is deeply rooted in our bodies, determined by bodily structure and modes of interaction with the

world, and carries profound evolutionary adaptive significance (Bobro, 1999; Yan, 2011; Yu et al., 2016; Zhao & Wu, 2019).

Container metaphor represents an ontological metaphor. Through the container boundary of our skin, we develop inside-outside sensorimotor experiences, which are further projected onto abstract conceptual domains to achieve understanding of complex conceptual meanings (Lakoff & Johnson, 1980). Previous research has demonstrated the objective reality of container metaphors in categorical cognition (Boot & Pecher, 2011) and self-cognition (Wang et al., 2019; Zhu & Zhang, 2001).

Do individuals utilize container space metaphor to understand and construct moral cognition? Although this linguistic metaphor is rarely used in Chinese, this does not imply the absence of psychological reality for this metaphor in cognition (Ye, 2017). From an evolutionary psychology perspective, seeking shelter is an unavoidable survival issue. Our ancient ancestors lived in a dangerous and unpredictable world. To avoid attacks from large predators and other potential dangers, they sought solid and safe shelters (Buss, 1999/2015). Consequently, the inside of a container may evoke sensorimotor experiences of safety and reliability. Regarding the world outside the container, our ancestors likely perceived it as dangerous and negative. This emotional perception of the external world represents an important adaptive mechanism that could be easily preserved through genes and might influence the metaphorical process of moral concepts through emotional valence (Li et al., 2019; Lu, Jia, et al., 2017).

Furthermore, Lakoff and Johnson (1999) argue that moral behavior is viewed as bounded activity that follows certain norms within permitted areas, whereas immoral behavior is seen as activity outside the rules that violates established norms and boundaries. Based on this moral boundary metaphor, the inside of a container is normative and reasonable and may thus be metaphorically mapped to morality, while the outside of a container is transgressive and "deviant," potentially mapped to immorality.

In summary, from these perspectives, the metaphorical association between moral concepts and container space has a reasonable basis. Accordingly, we propose the first research question: Does psychological reality exist for the metaphorical association between moral concepts and container space?

1.1.2 Mapping Bidirectionality and Theoretical Basis

Additionally, mapping directionality constitutes a fundamental issue in conceptual metaphor research. Some empirical studies suggest that metaphors are unidirectional (Casasanto & Boroditsky, 2008; Meier et al., 2004; Lu, Guo, & Jiang, 2017), while others support the bidirectionality view (Lu, Jia, et al., 2017; Jia & Jiang, 2016; Williams & Bargh, 2008; Zhong & Leonardelli, 2008). What causes this divergence? Theoretically, unidirectional support primarily comes from Conceptual Metaphor Theory (CMT), which was initially based on corpus analysis and represents speculative tracing of the universality and profundity of

metaphor existence, with limited generalization regarding its actual cognitive characteristics. Perceptual Symbols Theory (PST), grounded in cognitive evidence, better aligns with the patterns of metaphorical cognitive emergence and development. Therefore, the unidirectional mapping found in previous empirical research may result from retrieval failure rather than inherent unidirectionality. We hypothesize that the mapping of moral concepts onto container space should be bidirectional, representing the second research question of this study.

1.2.1 Metaphor Retrieval Hypothesis

Previous research indicates that metaphorical association formation conforms to the predictions of the CORE (CORrelations in Experience) principle, whereby different abstract concepts are shaped by sensorimotor experiences from different domains (Lakoff & Johnson, 1980; Lu, Jia, et al., 2017; Pitt & Casasanto, 2020). For moral concepts, direct sensorimotor experiences, robust evolutionary adaptive mechanisms, and more complex cultural and linguistic factors can all shape metaphorical associations, and these metaphorical mapping patterns may be simultaneously stored in long-term memory (Fischer, 2017; He et al., 2020; Holyoak & Stamenković, 2018), forming a hierarchical structure of mental metaphors (Casasanto & Bottini, 2014; Casasanto, 2017).

However, even when multiple metaphorical mapping patterns exist in cognition, this does not guarantee their retrieval. Many previous studies have conflated metaphor formation with metaphor retrieval. Failure to detect metaphorical congruency effects in experiments does not directly indicate the absence of metaphorical mapping in a particular direction; rather, it may result from metaphor retrieval failure. The retrieval process is often understood as the activation of mapping processes, which has been largely overlooked due to the influence of automatic metaphor mapping perspectives (He et al., 2020; Lu, Guo, et al., 2017; Yang et al., 2017). In fact, metaphor retrieval is independent of automaticity; it concerns whether metaphorical associations can be detected through behavioral measures, which is crucial for behavioral experimental results.

To emphasize the importance of the metaphor retrieval stage and based on existing research on the formation and storage patterns of metaphorical mapping, we propose the Metaphor Retrieval Hypothesis (Casasanto & Bottini, 2014; Casasanto, 2017; Fischer, 2017; Holyoak & Stamenković, 2018; Pitt & Casasanto, 2020). This hypothesis posits: (1) A metaphorical association may be shaped by various experiences, resulting in multiple hierarchically stored metaphorical mapping patterns that may be identical or completely opposite. The more frequently a metaphorical mapping pattern is used, the more easily it can be retrieved; (2) Metaphorical mapping is bidirectional, but due to differences in the sources and usage frequencies of different metaphorical mapping patterns, the mapping strength may be unbalanced in the two directions; (3) Different metaphorical associations vary in strength. Stronger metaphorical associations have higher retrieval priority, while weaker ones are less easily re-

trieved. Sensorimotor experiences, language, and culture collectively determine metaphorical association strength; (4) For weaker metaphorical associations, increasing processing depth and feature integration can both promote metaphorical association retrieval.

The moral concept-container space metaphor is relatively weak due to its scarcity in linguistic usage and is not easily retrieved. Therefore, according to the Metaphor Retrieval Hypothesis, enhancing processing depth and feature integration can both facilitate the retrieval of this metaphor.

1.2.2 Why Processing Depth Affects Metaphor Retrieval

When investigating moral concept metaphors, deeper encoding corresponds to stronger and more extensive neuronal activation in the brain, producing broader diffusion traces (Craik & Lockhart, 1972; Wang et al., 2020). Thus, the depth of conceptual or perceptual processing may influence the detection and retrieval of metaphorical associations.

Dual Coding Theory (DCT) of conceptual understanding posits that semantic and embodied factors function simultaneously in conceptual processing, but different factors dominate at different stages. In shallow conceptual processing, semantics plays the primary role, whereas deeper processing requires embodied sensorimotor experiences (Louwerse, 2008; Louwerse & Jeuniaux, 2010). Research has shown that conceptual processing depth indeed affects metaphorical associations (Liu et al., 2016; Wang et al., 2020). On the other hand, perceptual processing depth also influences metaphorical associations.

For instance, Liu et al. (2016) found that under deeper perceptual processing conditions, emotional faces and emotional concepts could influence each other, whereas this influence disappeared under shallow perceptual processing conditions, demonstrating that perceptual processing depth interferes with the retrieval of mapping. Huang et al. (2018) discovered that in a valence-brightness Stroop task, no metaphorical association emerged between emotional concepts and brightness perception, but this association was found when a go/no-go operation was added to the Stroop paradigm (Experiment 3), suggesting that pre-activation depth of non-judgment domain information affects metaphorical associations. Individuals process perceptual information quickly, which is often easily overlooked.

Compared to other perceptual experiences such as verticality, color, or size, container space perceptual information is more implicit among numerous features, with lower retrieval priority and requiring deeper perceptual processing. Therefore, the classic Stroop paradigm may have failed to retrieve the moral concept-container metaphor due to overly shallow perceptual processing depth. Enhancing the perceptual processing depth of container space for further investigation is a reasonable approach. Accordingly, this study proposes the third research question: Does perceptual processing depth affect the retrieval of the metaphorical association between moral concepts and container space?

1.2.3 Why Feature Integration Affects Metaphor Retrieval

According to Perceptual Symbols Theory, cognitive representation is closely related to and mutually influences the perceptual system. Thus, perceptual processing is intimately connected to the mental representation of abstract concepts. Individuals construct moral concepts through the diagramming and memory of perception. Consequently, whether the metaphorical association between moral concepts and container space can be effectively retrieved depends partly on whether container space perceptual features are adequately integrated into the cognitive representation of moral concepts (Barsalou, 1999). This process is termed feature integration.

Furthermore, Theory of Event Coding (TEC) posits that the encoding and storage of perceptual representations underlying perception and motor representations underlying action plans are not separate but exist in a common representational medium (Hommel et al., 2001). Amer et al. (2017) extended this to conceptual processing, arguing that the influence of prior concepts on subsequent concept processing depends on whether the features of the former and latter are bound into a common representation. Concepts generate a short-term unified representation by integrating all activated relevant event features, thereby creating connections among distributed features of events (Amer et al., 2017; Hommel, 1998, 2004; Kahneman et al., 1992; Treisman, 1998). If the two component domains of a metaphor are already connected, higher feature integration degree makes the metaphorical mapping pattern between domains more easily retrievable.

Based on these perspectives, feature integration degree may influence the retrieval of the metaphorical association between moral concepts and container space. We conducted experiments to verify this, representing the fourth research question of this study.

1.3 Research Framework

To adequately address the four research questions and provide evidence for the Metaphor Retrieval Hypothesis, this study conducted three experiments. Experiment 1, referencing Wang et al. (2018)' s spatial Stroop paradigm (Experiment 1), preliminarily examined the psychological reality of the metaphorical association between moral concepts and container space. Following the principle of economy, only the moral word judgment direction was designed for the Stroop task. Experiments 2a and 2b referenced the priming paradigm used in Wang et al. (2018), both employing two blocks with deeper and shallower processing depths to examine the bidirectional mapping of moral concept-container space metaphors and its susceptibility to perceptual processing depth. Finally, Experiments 3a and 3b built upon Experiment 1 by adding a go/no-go-like task requirement. By varying button responses, participants had to shift partial attention to container space perceptual features (or moral attribute features) while making moral attribute judgments (or inside-outside perceptual

information judgments). This preserved deeper perceptual processing depth in the Stroop task while further enhancing feature integration degree through the Stroop paradigm' s inherent characteristics and attentional investment, thereby effectively addressing the second and fourth research questions. The research framework is illustrated in Figure 1 [Figure 1: see original paper].

2.1 Purpose

This experiment adopted the spatial Stroop task from Wang et al. (2018) to investigate whether a metaphorical mapping exists from the target domain to the source domain between moral concepts and container space. If such mapping exists, participants should respond faster when judging moral words inside the container or immoral words outside the container.

2.2.1 Participants

Thirty right-handed participants (18 male, 12 female) with a mean age of 20.87 years (SD = 2.19) took part in the experiment. All reported no reading disabilities and had normal or corrected-to-normal vision. Participants received compensation upon completion.

2.2.2 Materials

Moral and immoral word stimuli were selected from Wang et al. (2020), which successfully detected metaphorical associations between moral concepts and horizontal orientation. These materials had demonstrated validity and thus met experimental requirements.

The experiment employed a 2 (word type: moral/immoral) × 2 (container space: inside/outside) within-subjects design. The dependent variable was reaction time in lexical judgment, with accuracy serving as a data screening index.

2.4 Procedure

The experiment used a spatial Stroop paradigm programmed in E-prime 2.0, following Wang et al. (2018). To avoid effects of word position, all words appeared at the screen center while circles appeared randomly at eight positions. For four positions, target words appeared inside the circle; for the other four, words appeared outside. Each position occurred equally often. Participants sat approximately 50 cm from a 14-inch IPS monitor (60Hz refresh rate, 1980$×$1080 resolution). After reading instructions, each trial began with a 500 ms red fixation cross "+" , followed by the target word at the same position, with a circle appearing randomly at different locations. Participants judged words as quickly and accurately as possible, pressing "F" for moral words and "J" for immoral words (with button assignments counterbalanced across approximately half the participants). After response or 3000 ms without response, a 500 ms blank screen appeared. The procedure is illustrated in Figure 2 [Figure 2: see

original paper]. Each of the 40 words appeared at all 8 circle positions, yielding 320 trials in the formal experiment, with rest intervals after every 80 trials.

Before the formal experiment, participants completed at least 16 practice trials using materials not included in the formal experiment. Practice trials provided feedback on accuracy and reaction time. After practice, participants confirmed understanding of requirements by pressing "Q" to proceed to the formal experiment or "P" to repeat practice.

2.5 Results

Thirty participants completed Experiment 1. Data from one participant with accuracy below 80% were excluded, leaving 29 participants for analysis. Error responses and extreme data points exceeding 2.5 standard deviations from the mean (697 trials, 7.5% of total data) were removed. Data were processed using SPSS 25.0, reporting both participant-based (F1) and item-based (F2) ANOVAs. As both analyses yielded similar descriptive results, only participant-based descriptive statistics are reported in tables for brevity, along with accuracy ANOVA results (consistent across all experiments).

Table 1 Mean Reaction Times and Standard Deviations for Moral Word Judgments (Participant Analysis)

Word Type	Inside Container	Outside Container
Moral	629.42$±74.04\|632.65±68.71\|\|$Immoral$|656.88±80.22|656.59±$83.46	

A 2$×2$repeatedmeasuresANOVAonreactiontimesrevealedasignificantmaineffectofwordtype$, F1(1, 28) = 19.86, p < 0.001, ηp^{2}$ = 0.42; F2(1, 316) = 30.98, p < 0.001, p^2 = 0.09, indicating faster judgments for moral than immoral words. The main effect of container space (inside/outside) was not significant, F1(1, 28) = 0.23, p = 0.633; F2(1, 316) = 0.21, p = 0.647. The word type × container space interaction was not significant, F1(1, 28) = 0.36, p = 0.553; F2(1, 316) = 0.08, p = 0.773.

Table 2 Mean Accuracy Rates and Standard Deviations for Moral Word Judgments (Participant Analysis)

Word Type	Inside Container	Outside Container
Moral	0.963$±0.040\|0.959±0.035\|\|$Immoral$|0.941±0.036|0.950±$0.035	

A 2$×2$repeatedmeasuresANOVAonaccuracyrevealedasignificantmaineffectofwordtype$, F(1, 28) = 5.91, p = 0.022, ηp^{2}$ = 0.17, indicating higher accuracy for moral than immoral words (M = 0.961 vs. M = 0.945). The main effect of container space was not significant, F(1, 28) = 0.63, p = 0.435. The word type × container space interaction was not significant, F(1, 28) = 2.22, p = 0.147.

2.6 Discussion

The results showed faster and more accurate responses to moral words than immoral words, consistent with previous findings (Lu, Guo, et al., 2017; Meier et al., 2007; Yin et al., 2014; Yang et al., 2017). This may reflect an attentional bias toward immorality or result from individual motivational approach-avoidance systems (Liu et al., 2016; Yang et al., 2017).

However, no significant metaphorical congruency effect was observed, possibly due to metaphor retrieval failure. First, container space perceptual information represents categorical perception. Compared to vertical spatial or color perception, it lacks temporal advantage among numerous feature attributes and is easily masked. Second, the Stroop task only required moral word judgments while ignoring spatial relationships between circles and words, leading to shallow processing of container space perceptual information. Most importantly, Chinese rarely employs container metaphors for moral concepts, resulting in weak linguistic reality and consequently weak metaphorical mapping patterns. Therefore, we cannot conclude that the moral concept-container metaphor does not exist. Both weak metaphorical association and shallow processing depth may have contributed to the null results. Experiments 2 and 3 subsequently adopted a priming paradigm to enhance perceptual processing depth for further investigation of bidirectional metaphorical associations.

3 Experiment 2a: Influence of Perceptual Processing Depth on Target-to-Source Mapping in Moral Concept-Container Metaphor

Experiment 2a employed the priming paradigm from Wang et al. (2018) to investigate whether individuals exhibit target-to-source mapping in moral concept-container space metaphorical understanding under deeper and shallower container space perceptual priming conditions. If such mapping exists, container-inside priming should facilitate moral word processing relative to immoral words, while container-outside priming should facilitate immoral word processing, at least in the deeper processing block.

3.2.1 Participants

Thirty right-handed participants (10 male, 20 female) with a mean age of 22.60 years (SD = 1.99) participated. All reported no reading disabilities and had normal or corrected-to-normal vision. Participants received compensation upon completion.

3.2.2 Materials

Ten moral and ten immoral words were selected from materials used in Experiment 1. Moral words (M = 7.25) differed significantly from the midpoint of 5 in a one-sample t-test, $t(9) = 13.60$, $p < 0.001$. Immoral words (M = 2.12) also

differed significantly from the midpoint, t(9) = -38.78, p < 0.001, confirming distinct valences. Word frequency analysis showed no significant difference between moral (M = 0.0149) and immoral words (M = 0.0141), t(18) = 0.35, p > 0.05. Stroke count evaluation revealed no significant difference between moral (M = 18.90, SD = 3.21) and immoral words (M = 18.80, SD = 3.45), t(18) = 0.08, p > 0.05, confirming adequate stimulus control.

The experiment used a 2 (processing depth: deep/shallow) × 2 (word type: moral/immoral) × 2 (container space: inside/outside) within-subjects design. Reaction time in lexical judgment served as the dependent variable, with accuracy as a data screening index.

3.4 Procedure

The experiment employed a priming paradigm programmed in E-prime 2.0, following Wang et al. (2018). Participants sat approximately 50 cm from a 15.6-inch LED monitor (60Hz refresh rate, 1366$×$768 resolution). The experiment consisted of two blocks, each preceded by instructions. Each trial began with a 500 ms red fixation cross "+" , followed by a circle (approximately 40% of screen width) appearing 20% left or right of center, with its left or right boundary tangent to the screen' s central vertical line. After 200 ms, letters "p" or "q" appeared randomly at positions 20% left or right of center. Different blocks required different judgment tasks. In one block, participants performed categorical judgments, pressing "P" for "p" and "Q" for "q" . In the other block, they performed location judgments, pressing "E" for inside the circle and "I" for outside. Target words then appeared at the letter' s position, and participants made moral judgments as quickly and accurately as possible, pressing "F" for moral words and "J" for immoral words. After response or 3000 ms without response, a 500 ms blank screen appeared. Button assignments and block order were counterbalanced across participants. The procedure is illustrated in Figure 3 [Figure 3: see original paper]. Each block contained 160 trials, with each word appearing 8 times, totaling 320 trials. Practice trials preceded the formal experiment, following the same standard as Experiment 1.

3.5 Results

Figure 3 [Figure 3: see original paper] Experiment 2a Procedure

Thirty participants completed Experiment 2a. Data from one participant with letter judgment accuracy below 80% in one block were excluded, leaving 29 participants for analysis. Extreme data points (780 trials, 8.4% of total data) were removed using the same method as Experiment 1. Data were processed using SPSS 25.0. As the primary research focus was metaphorical mapping under different processing depths, separate 2$×$2 repeated measures ANOVAs were conducted for deep and shallow processing conditions.

Table 3 Mean Reaction Times and Standard Deviations for Moral Word Judg-

ments (Participant Analysis)

Processing Depth	Word Type	Inside Container	Outside Container
Deep	Moral	656.87$±116.66|708.85±103.84||*Deep*|*Immoral*|711.77±134.39|729.15±14	

Under deep processing conditions, ANOVA revealed a significant main effect of word type, F1(1, 28) = 17.68, p < 0.001, p^2 = 0.39; F2(1, 156) = 40.67, p < 0.001, p^2 = 0.21, indicating faster judgments for moral than immoral words. The main effect of container space was significant, F1(1, 28) = 11.86, p = 0.002, p^2 = 0.30; F2(1, 156) = 14.60, p < 0.001, p^2 = 0.09, indicating faster judgments following inside versus outside priming. The word type × container space interaction was marginally significant, F1(1, 28) = 2.92, p = 0.099, p^2 = 0.09; F2(1, 156) = 4.20, p = 0.042, p^2 = 0.03. Simple effects analysis showed that for moral words, the inside-outside difference was significant, F1(1, 28) = 10.31, p = 0.003, p^2 = 0.27; F2(1, 156) = 13.75, p < 0.001, Cohen' s d = -0.34, with faster responses to moral words following inside versus outside priming. For immoral words, the inside-outside difference was not significant, F1(1, 28) = 1.39, p = 0.25; F2(1, 156) = 1.25, p = 0.264.

Under shallow processing conditions, the main effect of word type was significant, F1(1, 28) = 9.08, p = 0.005, p^2 = 0.25; F2(1, 156) = 13.96, p < 0.001, p^2 = 0.08. The main effect of container space was not significant, F1(1, 28) = 0.01, p = 0.928; F2(1, 156) = 0.001, p = 0.971. The word type × container space interaction was not significant, F1(1, 28) = 0.51, p = 0.482; F2(1, 156) = 0.21, p = 0.649.

Table 4 Mean Accuracy Rates and Standard Deviations for Moral Word Judgments (Participant Analysis)

Processing Depth	Word Type	Inside Container	Outside Container
Deep	Moral	0.973$±0.031|0.962±0.035||*Deep*|*Immoral*|0.961±0.054|0.960±0.047||*Sha*	

Accuracy ANOVA under deep processing showed a marginally significant main effect of word type, F(1, 28) = 3.52, p = 0.071, p^2 = 0.11. The main effect of container space was not significant, F(1, 28) = 0.14, p = 0.710. The interaction was not significant, F(1, 28) = 0.02, p = 0.904. Under shallow processing, no significant effects were found (all ps > 0.20).

3.6 Discussion

Experiment 2a results showed significant main effects of word type under both processing depths, consistent with Experiment 1. More importantly, under deep processing, a metaphorical association emerged between moral concepts and container space. Moral words were judged significantly faster following inside

versus outside priming, while no such difference appeared for immoral words. Under shallow processing, no metaphorical mapping was detected. These results partially support our hypothesis that processing depth affects moral concept-container metaphor retrieval. However, Experiment 2a only revealed the moral-inside association, with the immoral-outside association remaining unclear. We therefore conducted Experiment 2b.

4 Experiment 2b: Influence of Perceptual Processing Depth on Source-to-Target Mapping in Moral Concept-Container Metaphor

Experiment 2b also used a priming paradigm to investigate whether individuals exhibit source-to-target mapping in moral concept-container space metaphorical understanding under deeper and shallower container space perceptual conditions. If such mapping exists, moral word priming should facilitate letter processing inside the container, while immoral word priming should facilitate letter processing outside the container, at least in the deeper processing block.

4.2.1 Participants

Thirty right-handed participants (11 male, 19 female) with a mean age of 19.27 years (SD = 2.20) participated. All reported no reading disabilities and had normal or corrected-to-normal vision. Participants received compensation upon completion.

4.3 Procedure

The design and procedure mirrored Experiment 2a, except participants performed moral word judgment first, followed by letter judgment. The procedure is illustrated in Figure 4 [Figure 4: see original paper]. Practice trials preceded the formal experiment, following the same standard as Experiment 1.

4.4 Results

Figure 4 [Figure 4: see original paper] Experiment 2b Procedure

All thirty participants met experimental requirements, so all data were included in analysis. Extreme data points (843 trials, 8.8% of total data) were removed using the same method as Experiment 2. Data were processed using SPSS 25.0, with analysis methods similar to Experiment 2.

Table 5 Mean Reaction Times and Standard Deviations for Letter Judgments (Participant Analysis)

Processing Depth	Word Type	Inside Container	Outside Container
Deep	Moral	609.61$±146.76|648.89±148.33||*Deep*|*Immoral*|610.06±143.87|610.00±15	

Under deep processing conditions, ANOVA revealed a significant main effect of word type, F1(1, 29) = 11.44, p = 0.002, p^2 = 0.28; F2(1, 156) = 13.90, p < 0.001, p^2 = 0.08, indicating faster letter judgments following moral versus immoral priming. The main effect of container space was not significant, F1(1, 29) = 0.01, p = 0.905; F2(1, 156) = 0.03, p = 0.855. The word type × container space interaction was marginally significant, F1(1, 29) = 3.08, p = 0.090, p^2 = 0.10; F2(1, 156) = 7.54, p = 0.007, p^2 = 0.05. Simple effects analysis showed that inside the container, the difference between moral and immoral priming was significant, F1(1, 29) = 9.89, p = 0.004, p^2 = 0.25; F2(1, 156) = 21.08, p < 0.001, Cohen' s d = -1.03, with faster letter judgments following moral versus immoral priming. Outside the container, this difference was not significant, F1(1, 29) = 0.18, p = 0.673; F2(1, 156) = 0.49, p = 0.486.

Under shallow processing conditions, the main effect of word type was not significant, F1(1, 29) = 0.001, p = 0.978; F2(1, 156) = 0.06, p = 0.806. The main effect of container space was significant, F1(1, 29) = 46.34, p < 0.001, p^2 = 0.62; F2(1, 156) = 31.35, p < 0.001, p^2 = 0.17, indicating faster judgments for letters inside versus outside circles. The interaction was not significant, F1(1, 29) = 0.001, p = 0.980; F2(1, 156) = 0.002, p = 0.965.

Table 6 Mean Accuracy Rates and Standard Deviations for Letter Judgments (Participant Analysis)

Processing Depth	Word Type	Inside Container	Outside Container
Deep	Moral	0.953$±0.054|0.943±0.057||*Deep*|*Immoral*|0.974±0.032|0.973±0.023||*Sha*	

Under deep processing conditions, accuracy ANOVA showed a significant main effect of word type, F(1, 29) = 10.01, p = 0.004, p^2 = 0.26. The main effect of container space was not significant, F(1, 29) = 0.03, p = 0.865. The interaction was marginally significant, F(1, 29) = 3.21, p = 0.084. Simple effects analysis revealed no significant effects for either moral or immoral priming conditions (ps > 0.10). Under shallow processing, no significant effects were found (ps > 0.13).

4.5 Discussion

Experiment 2b results showed a significant interaction between word type and container space under deep perceptual processing, with faster letter judgments inside the container following moral versus immoral priming, but no difference outside the container. Under shallow processing, the interaction was not significant. Consistent with Experiment 2a, only the moral-inside association was found, with the immoral-outside association remaining unclear. Although Experiments 2a and 2b demonstrated metaphorical associations from both mapping directions under deep perceptual processing, these associations appeared

unstable even under deep processing. Therefore, it was necessary to further enhance feature integration degree while maintaining deep perceptual processing, leading to Experiment 3.

5 Experiment 3a: Influence of Feature Integration Degree on Target-to-Source Mapping in Moral Concept-Container Metaphor

Experiment 3a used a Stroop paradigm with added task demands to investigate target-to-source mapping in moral concept-container space metaphorical understanding under higher feature integration. If such mapping exists, participants should respond faster to moral words inside containers and immoral words outside containers.

5.2.1 Participants

Thirty right-handed participants (12 male, 18 female) with a mean age of 20.03 years (SD = 2.53) participated. All reported no reading disabilities and had normal or corrected-to-normal vision. Participants received compensation upon completion.

5.3 Procedure

The experiment followed the same design as Experiment 1 but used a 15.6-inch LED monitor (60Hz refresh rate, 1366$×$768 resolution). Participants judged target words as quickly and accurately as possible using different buttons based on location: for words inside the circle, "F" for moral and "J" for immoral; for words outside the circle, "D" for moral and "K" for immoral. Button assignments were counterbalanced across participants. The procedure is illustrated in Figure 5 [Figure 5: see original paper]. Practice trials preceded the formal experiment, following the same standard as Experiment 1.

5.4 Results

Figure 5 [Figure 5: see original paper] Experiment 3a Procedure

Thirty participants completed Experiment 3a. Data from three participants with accuracy below 70% were excluded, leaving 27 participants for analysis. Extreme data points beyond 3 standard deviations (900 trials, 10.4% of total data) were removed. Data were processed using SPSS 25.0.

Table 7 Mean Reaction Times and Standard Deviations for Word Judgments (Participant Analysis)

Word Type	Inside Container	Outside Container
Moral	902.36$±173.14\|1007.98±226.86\|\|$Immoral$|1009.14±204.53|984.90±$207.61	

A 2$×2repeatedmeasuresANOVArevealedasignificantmaineffectofwordtype, $F1(1, 26) = 8.58, p = 0.007, \eta p\hat{}\{2\}$ = 0.25; F2(1, 316) = 28.11, p < 0.001, p^2 = 0.08, indicating faster judgments for moral than immoral words. The main effect of container space was significant, F1(1, 26) = 14.27, p = 0.001, p^2 = 0.35; F2(1, 316) = 18.30, p < 0.001, p^2 = 0.06, indicating faster judgments for words inside versus outside circles. The word type × container space interaction was significant, F1(1, 26) = 31.27, p < 0.001, p^2 = 0.55; F2(1, 316) = 43.50, p < 0.001, p^2 = 0.12. Simple effects analysis showed that for moral words, the inside-outside difference was significant, F1(1, 26) = 30.40, p < 0.001, p^2 = 0.54; F2(1, 316) = 54.45, p < 0.001, Cohen' s d = -1.26, with faster responses to moral words inside versus outside containers. For immoral words, the difference was also significant in participant analysis, F1(1, 26) = 4.36, p = 0.047, p^2 = 0.14, though not in item analysis, F2(1, 316) = 2.48, p = 0.117, with faster responses to immoral words outside versus inside containers.

Table 8 Mean Accuracy Rates and Standard Deviations for Word Judgments (Participant Analysis)

Word Type	Inside Container	Outside Container
Moral	0.948$±0.042\|0.879±0.096\|\|$Immoral$|0.881±0.070|0.931±$0.045	

Accuracy ANOVA revealed no significant main effect of word type, F(1, 26) = 0.47, p = 0.500. The main effect of container space was not significant, F(1, 26) = 0.91, p = 0.349. The interaction was significant, F(1, 26) = 21.01, p < 0.001, p^2 = 0.45. Simple effects analysis showed that for moral words, accuracy was higher inside versus outside containers, F(1, 26) = 14.67, p = 0.001, p^2 = 0.36. For immoral words, accuracy was higher outside versus inside containers, F(1, 26) = 11.13, p = 0.003, p^2 = 0.30.

5.5 Discussion

Experiment 3a results showed a significant interaction between word type and container space, with faster responses to moral words inside containers and immoral words outside containers. Compared to Experiment 2a' s deep processing condition, the interaction was more significant (Experiment 2a: p = 0.099; Experiment 3a: p < 0.001) with substantially larger effect size (Experiment 2a: p^2 = 0.09; Experiment 3a: p^2 = 0.55). Moreover, Experiment 3a detected the immoral-outside metaphorical congruency effect that Experiment 2a failed to find. These results demonstrate that feature integration effectively promotes the detection of moral concept-container space metaphorical associations.

6 Experiment 3b: Influence of Feature Integration Degree on Source-to-Target Mapping in Moral Concept-Container Metaphor

Building on Experiment 3a, Experiment 3b also used a Stroop paradigm with added task demands to investigate source-to-target mapping in moral concept-container space metaphorical understanding under higher feature integration. If such mapping exists, participants should respond faster to moral words inside containers and immoral words outside containers. Comparing the significance and effect sizes of metaphorical congruency effects between Experiments 3b and 2b can further illuminate how feature integration degree influences metaphorical mapping in this direction.

6.2.1 Participants

Thirty right-handed participants (12 male, 18 female) with a mean age of 19.57 years (SD = 2.46) participated. All reported no reading disabilities and had normal or corrected-to-normal vision. Participants received a gift upon completion.

6.3 Procedure

The design and procedure were similar to Experiment 3a, with the task changed to spatial relationship judgment. Specifically, for moral words, participants pressed "F" if inside the circle and "J" if outside; for immoral words, they pressed "D" if inside and "K" if outside. The procedure is illustrated in Figure 6 [Figure 6: see original paper].

6.4 Results

Figure 6 [Figure 6: see original paper] Experiment 3b Procedure

Thirty participants completed Experiment 3b. Data from three participants with accuracy below 70% were excluded, leaving 27 participants for analysis. Extreme data points (862 trials, 10.0% of total data) were removed using the same method as Experiment 3a. Data were processed using SPSS 25.0.

Table 9 Mean Reaction Times and Standard Deviations for Container Space Judgments (Participant Analysis)

Word Type	Inside Container	Outside Container
Moral	981.94$±151.04\|1142.86±176.78\|\|$Immoral$|1142.55±179.44|1080.46±$175.93	

A 2$×2 repeated measures ANOVA revealed a significant main effect of word type, $F1(1, 26) = 23.24, p < 0.001, \eta p^{\wedge}\{2\}$ = 0.47; F2(1, 316) = 26.64, p < 0.001, p^2 = 0.08, indicating faster judgments for moral than immoral words. The main effect of

container space was significant, F1(1, 26) = 8.32, p = 0.008, p^2 = 0.24; F2(1, 316) = 24.65, p < 0.001, p^2 = 0.07, indicating faster judgments for words inside versus outside containers. The word type × container space interaction was significant, F1(1, 26) = 91.39, p < 0.001, p^2 = 0.78; F2(1, 316) = 132.99, p < 0.001, p^2 = 0.30. Simple effects analysis showed that for moral words, responses were faster inside versus outside containers, F1(1, 26) = 61.39, p < 0.001, p^2 = 0.70; F2(1, 316) = 125.89, p < 0.001, Cohen' s d = -1.90. For immoral words, responses were faster outside versus inside containers, F1(1, 26) = 9.02, p = 0.006, p^2 = 0.26; F2(1, 316) = 19.95, p < 0.001, Cohen' s d = 0.73.

Table 10 Mean Accuracy Rates and Standard Deviations for Container Space Judgments (Participant Analysis)

Word Type	Inside Container	Outside Container
Moral	0.965$±0.032\|0.907±0.057\|\|$Immoral$|0.888±0.055|0.929±$0.056	

Accuracy ANOVA revealed a significant main effect of word type, F(1, 26) = 13.73, p = 0.001, p^2 = 0.36. The main effect of container space was not significant, F(1, 26) = 0.98, p = 0.331. The interaction was significant, F(1, 26) = 48.64, p < 0.001, p^2 = 0.67. Simple effects analysis showed that for moral words, accuracy was higher inside versus outside containers, F(1, 26) = 25.30, p < 0.001, p^2 = 0.51. For immoral words, accuracy was higher outside versus inside containers, F(1, 26) = 12.58, p = 0.002, p^2 = 0.34.

6.5 Discussion

Experiment 3b results showed a significant interaction between word type and container space, with faster responses when moral words appeared inside containers and immoral words appeared outside containers. Compared to Experiment 2b' s deep processing condition, the interaction was more significant (Experiment 2b: p = 0.090; Experiment 3b: p < 0.001) with substantially larger effect size (Experiment 2b: p^2 = 0.10; Experiment 3b: p^2 = 0.78). Moreover, Experiment 3b detected the immoral-outside metaphorical congruency effect that Experiment 2b failed to find.

7 General Discussion

This study investigated the psychological reality and bidirectional mapping of moral concept-container space metaphorical associations, as well as the influence of processing depth and feature integration on metaphor retrieval, through three experiments. Experiment 1 using the classic spatial Stroop paradigm found no metaphorical association. Experiment 2 using the priming paradigm revealed psychological reality and bidirectional mapping under deeper perceptual processing, though the association remained incomplete. Experiment 3 using the

Stroop paradigm with added task demands demonstrated more complete psychological reality and more pronounced bidirectional mapping under higher feature integration.

7.1 The Psychological Reality of Moral Concept-Container Space Metaphor

Experiment 1's spatial Stroop task failed to detect metaphorical association, possibly because Chinese rarely uses inside-outside expressions for moral attributes, preventing repeated activation and consolidation in daily life and resulting in weak psychological reality (Barsalou, 1999; Lee & Schwarz, 2012; Wang et al., 2020). Considering retrieval failure possibilities, Experiments 2 and 3 successfully detected this metaphor' s psychological reality from different angles.

Through interaction with the external world, humans develop perceptual experiences of inside-outside distinctions and project them onto other thinking domains, forming an implicit internal-external spatial schema in the mind. This categorical distinction helps understand and represent the complex external world and may be further structured onto meaning construction of abstract concepts like morality (Williams et al., 2009). From an evolutionary psychology perspective, inside containers corresponds to habitats representing warmth and safety, while outside corresponds to nature representing unknown and danger. This inside-outside container perceptual experience is crucial for species survival and reproduction, carrying important evolutionary adaptive significance (Buss, 1999/2015). Additionally, Lakoff et al. (1999) argue that morality is generally viewed as bounded behavior following certain norms. In traditional Chinese culture, Confucian "rules" deeply embody this moral boundary metaphor. The Analects states: "The gentleman guides with words and restrains with deeds. Thus, in speech he considers the outcome; in action he examines the consequences. Therefore, people are cautious in speech and careful in conduct," reflecting that a cultivated, moral "gentleman" should possess self-restraint and inner discipline, demonstrating characteristics of moral container metaphor.

7.2 Bidirectionality of Metaphorical Mapping

Although some scholars argue that metaphors are unidirectional (Lakoff & Johnson, 1980), this conclusion may derive primarily from linguistic evidence analysis, limiting its applicability and inadequately explaining the psychological reality of metaphor generation. In early life, concrete concept sensorimotor experiences are easily understood, and individuals naturally use these concrete concepts to acquire abstract concepts (Lee & Schwarz, 2012). Once metaphorical associations form, they establish neural connections in the brain. Through habitual projection based on sensorimotor experience, source and target domains develop potential mechanisms for cross-domain neural resonance, representing the internal basis for metaphorical association' s psychological reality (Barsalou, 1999). In this sense, metaphorical associations are bidirectional. Some

studies' failure to detect bidirectional mapping may result from retrieval failure, which could be overcome by increasing processing depth or enhancing feature integration (Huang et al., 2018; Wang et al., 2020).

7.3 The Influence of Processing Depth and Feature Integration on Metaphor Retrieval

Experiment 2 manipulated container space perceptual processing depth and found that depth indeed affected metaphorical association and mapping bidirectionality. From the levels-of-processing perspective, different processing depths of objects lead to differences in memory effects and meaning understanding (Craik & Lockhart, 1972). More specifically, different processing depths imply differences in neural activation scope for corresponding features. According to Perceptual Symbols Theory, greater activation of perceptual information makes it easier to evoke corresponding perceptual schemas and project them onto metaphorical frameworks (Barsalou, 1999).

Additionally, Huang et al. (2018)' s activation hypothesis can explain these results, positing that concrete concept processing is faster than abstract concept processing, and that non-judgment domain processing depth affects metaphorical congruency effects. More importantly, our results support the Metaphor Retrieval Hypothesis. Some previous studies detected metaphorical associations without deep perceptual processing, suggesting that different metaphorical associations vary in strength (Huang et al., 2018; Yang et al., 2017). Different perceptual information shapes moral concepts through different mechanisms. Container metaphor involves categorical relationships between two objects, which are less automatically processed. Without conscious attention, shallow processing means insufficient cognitive resources are allocated, making it susceptible to masking by higher-priority perceptual features (Lu, Guo, et al., 2017). Moreover, the weak linguistic reality foundation of moral concept-container metaphor places this mapping pattern at a lower retrieval priority. Enhancing processing depth awakens greater neural activation, potentially highlighting this weak metaphorical association and enabling its retrieval.

Although Experiment 2 detected metaphorical association and bidirectional mapping under deep perceptual processing, the interaction significance and effect sizes in participant-based ANOVA were modest (Experiment 2a: $p = 0.099$, $p^2 = 0.09$; Experiment 2b: $p = 0.090$, $p^2 = 0.10$), and the metaphorical associations were incomplete. Therefore, Experiment 3 further investigated moral concept-container metaphor under higher feature integration in both mapping directions. Results showed strong metaphorical congruency effects in both directions, with highly significant interactions and substantially increased effect sizes (Experiment 3a: $p < 0.001$, $p^2 = 0.55$; Experiment 3b: $p < 0.001$, $p^2 = 0.78$), demonstrating that feature integration promotes detection of moral concept-container metaphor.

These results align with Metaphor Retrieval Hypothesis predictions. Percep-

tual symbols are neural representations of the external world centered on perception, not isolated islands outside conceptual processing. Enhancing integration between container space perceptual features and moral concepts essentially involves simultaneous activation of source and target domains within limited cognitive resources, making previously stored metaphorical mapping patterns more accessible (Barsalou, 1999). Furthermore, according to Theory of Event Coding, feature integration facilitates binding of distributed attribute features into common representations (Amer et al., 2017; Hommel, 1998, 2004; Kahneman et al., 1992; Treisman, 1998). For stronger metaphorical associations, even low feature integration allows detection due to high retrieval priority. However, moral concept-container metaphor, as a weaker association, is particularly susceptible to feature integration degree, potentially leading to retrieval failure. Thus, feature integration does not affect metaphor formation or change association strength itself, but influences the retrieval process by enhancing mapping pattern priority, making it more detectable.

7.4 Metaphor Formation and Retrieval: A Dissociation Perspective

This series of experiments investigating moral concept-container space metaphor and its influencing factors provides strong support for the Metaphor Retrieval Hypothesis. Adopting a dissociative perspective that separates metaphor formation origins from retrieval processes has important theoretical significance and value, helping explain and clarify contradictions in existing research.

Previous metaphor studies often overlooked retrieval processes, leading to many contradictory results. First, the Metaphor Retrieval Hypothesis can well explain the pervasive contradiction between unidirectional and bidirectional mapping findings (Casasanto & Boroditsky, 2008; Lu, Guo, et al., 2017; Lu, Jia, et al., 2017; Jia & Jiang, 2016; Meier et al., 2004; Williams & Bargh, 2008; Zhong & Leonardelli, 2008). The discovered unidirectional mapping does not reflect inherent characteristics of metaphorical associations but rather results from partial mapping retrieval failure, stemming from the complexity of metaphorical mapping pattern origins and construction, influenced by experimental paradigms, participants' cultural-linguistic backgrounds, and metaphorical association strength.

Second, research shows that mental time line (MTL) and mental number line (MNL) directions are cross-culturally influenced, but the sensorimotor experiences affecting them differ. For MTL, reading and writing experience plays the primary role (Casasanto & Bottini, 2014; Pitt & Casasanto, 2020), whereas for MNL, finger counting direction is key (Pitt & Casasanto, 2020; Riello & Rusconi, 2011). This indicates that different abstract concepts contain different metaphorical mapping patterns, with different cultural-linguistic backgrounds producing different pattern strengths and retrieval priorities. The Metaphor Retrieval Hypothesis excellently explains these inconsistent cross-cultural findings.

Moreover, even for participants from the same cultural-linguistic background, consistent metaphor retrieval depends on retrieval processes. Examples include the duality problem in Chinese moral concept horizontal orientation metaphors (Yang et al., 2017; Wang et al., 2020) and the reverse MTL or MNL effects caused by immediate practice (Pitt & Casasanto, 2020). Overall, the duality problem of metaphorical association retrieval in moral, temporal, and numerical abstract concept domains can be explained by the Metaphor Retrieval Hypothesis: metaphorical connection representations store two completely opposite mapping patterns. Under normal circumstances, more commonly used patterns are more easily retrieved, while immediate training grants higher retrieval priority to the most relevant pattern, leading to reverse associations.

Thus, the Metaphor Retrieval Hypothesis separates metaphor formation from retrieval, clearly explaining the internal operating mechanisms of metaphorical associations. This not only helps clarify previous research controversies but also identifies processing depth and feature integration as two important factors promoting metaphor retrieval, which can be extended to other metaphor research domains, providing guidance and reference value for future studies.

8 Conclusions, Limitations, and Future Directions

This study draws several conclusions: (1) Psychological reality exists for the metaphorical association between moral concepts and container space, represented as moral inside and immoral outside, with bidirectional mapping; (2) Increasing perceptual processing depth or enhancing feature integration degree both promote retrieval of moral concept-container space metaphorical associations; (3) Metaphor formation and retrieval exhibit certain dissociation, with the Metaphor Retrieval Hypothesis providing good explanatory power.

However, this study has limitations. First, to avoid individual differences, Experiment 2 set deep and shallow processing as two within-subject blocks. To balance interference, block order was counterbalanced across participants and materials were reduced to decrease trials, but this created inequality in experimental materials between Experiments 2 and 1. Future research should ensure material consistency for better comparison. Second, while enhancing feature integration in Experiment 3, perceptual or moral feature processing depth may have been inadvertently increased, affecting conclusion validity. Future studies should consider balancing approximate reaction times across conditions to eliminate extraneous influences. Additionally, although data exclusion criteria were relaxed in Experiment 3, the inherently difficult task still resulted in relatively high data exclusion rates.

This study verified the dissociation between metaphor formation and retrieval, supporting the Metaphor Retrieval Hypothesis. Future research can further explore: First, previous studies that failed to detect or found incomplete metaphorical congruency effects, such as moral concepts with red-white color perception or upright/skew font perception (Yang et al., 2017), and studies finding unidirec-

tional mapping like moral concepts with size perception (Lu, Guo, et al., 2017), may all result from metaphor retrieval failure. Second, research shows immediate training can cause reverse MNL or MTL effects (Pitt & Casasanto, 2020). Does immediate training also affect moral concept metaphors? According to the Metaphor Retrieval Hypothesis, for metaphorical associations potentially containing two opposite mapping patterns, corresponding immediate training may cause reverse associations, warranting further investigation.

References

Amer, T., Gozli, D. G., & Pratt, J. (2017). Biasing spatial attention with semantic information: an event coding approach. *Psychological Research, 82*(3), 1–19.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577–609.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Buss, M. D. (2015). *Evolutionary Psychology: the New Science of the Mind* (Y. Zhang, & K. Jiang, Trans.). New York and London: Taylor & Francis Group. (Original work published 1999)

Bobro, M. E. (1999). Leibniz on embodiment and the moral order. *The Southern journal of philosophy, 37*(3), 377–

Boot, I., & Pecher, D. (2011). Representation of categories: metaphorical use of the container schema. *Experimental Psychology, 58*(2), 162–170.

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin, 143*(3), 263–292.

Casasanto, D. (2011). Different bodies, different minds: The body specificity of language and thought. *Current Directions in Psychological Science, 20*(6), 378–383.

Casasanto, D. (2017). The hierarchical structure of mental metaphors. In B. Hampe (Ed.), *Metaphor: Embodied cognition and discourse* (pp. 46–61). Cambridge, UK: Cambridge University Press.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition, 106*, 579–

Casasanto, D., & Bottini, R. (2014). Mirror reading can reverse the flow of time. *Journal of Experimental Psychology: General, 143*, 473–479.

Chen, X., Jiang, Q., Hou, M., & Zhu, M. Y. (2014). Embodied morality: A new approach in moral psychology. *Psychological Development and Education, 30*(6), 664–672.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Fischer, E. (2017). Two analogy strategies: The cases of mind metaphors and introspection. *Connection Science. 30*(2), 211-243

Gibbs, R. W., Jr. (2006). *Embodiment and cognitive science*. Cambridge, UK: Cambridge University Press.

He, X. L., Chen, J. (2020). Cognitive development of multiple metaphors of power concepts in 3∼5 year-old children. *Acta Psychologica Sinica, 52*(2):149-161.

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: a critical review of theories and evidence. *Psychological Bulletin, 144*(6), 641-671.

Hommel, B. (1998). Automatic stimulus-response translation in dual- task performance. *Journal of Experimental Psychology-Human Perception and Performance, 24*, 1368-1384.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494-500.

Hommel, B., Mu¨sseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849-878.

Huang, Y., Tse, C.-S., & Xie, J. (2018). The bidirectional congruency effect of brightness-valence metaphoric association in the Stroop-like and priming paradigms. *Acta Psychologica, 189*(10), 76-92.

Jia, L., &Jiang, G. F. (2016). The vertical spatial metaphor of moral concepts : Psychological reality and bidirectional mapping. *Psychological Development and Education, 32*(2), 158-165.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files –Object-specific integration of information. *Cognitive Psychology, 24*,175-219.

Lakoff, G., & Johnson, M. (Eds). (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (Eds). (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.

Lee, S. W. S., & Schwarz, N. (2012). Bidirectionality, mediation, and moderation of metaphorical effects: The embodiment of social suspicion and fishy smells. *Journal of Personality and Social Psychology, 103*(5), 737-

Li, Y., Zhang, C., & Wang, R. (2019). The effect of moral emotions on the metaphorical mapping of morality and its neural mechanism. *Advances in Psychological Science, 27*(7), 1224-1231.

Liu, W. J., Shen, M. Q., Li, Y., &Wang, R. M. (2016). The interaction between emotional concept processing and emotional face perception. *Acta Psychologica Sinica,48*(2), 163-173.

Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review, 15*, 838-

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition, 114*, 96-104.

Lu, Z. Y., Guo, S. P., &Jiang, Z. L. (2017). The size metaphor of moral concepts: Psychological reality and mapping relationship. *Journal of South China Normal University (Social Science Edition), 49*(2),70-78.

Lu, Z. Y., Jia, L. L., &Zhai, D. X. (2017). The mapping for vertical spatial metaphor of the moral concepts: Bidirectional and unbalanced. *ActaPsychologica Sinica,49*(2), 186-196.

Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white: Automatic inferences about stimulus valence based on brightness. *Psychological Science, 15*,82-87.

Meier, B. P., Sellbom, M., & Wygant, D. B. (2007). Failing to take the moral high ground: Psychopathy and the vertical representation of morality. *Personality and Individual Differences, 43*(4), 757-767.

Pitt, B., & Casasanto, D. (2020). The correlations in experience principle: How culture shapes concepts of time and number. *Journal of Experimental Psychology: General, 149*(6),1048-1070.

Riello, M., & Rusconi, E. (2011). Unimanual SNARC effect: Hand matters. *Frontiers in Psychology, 2*, 372.

Treisman, A. M. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 353*, 1295-1306.

Wang, X. X., Jiang, S., Zhang, J. J. (2018). Effect of the spatial linguistic symbol on the container metaphor of seniority rules. *ActaPsychologica Sinica, 50*(09), 29-40.

Wang, B., Fu, Y., Zhang, J. J. (2019). Influence of language and culture on retrieval-induced forgetting under the self-referential condition: Evidence from the Han and the Mosuo. *ActaPsychologica Sinica, 51*(04), 60-71.

Wang, C. X., Ma, J. P., Deng, J., Yang, Z. W., Ye, Y. D. (2020). The depth of conceptual processing affects the metaphorical connection between moral concepts and horizontal orientation. *ActaPsychologica Sinica, 52*(4), 426-439.

Wang, Z., & Lu, Z. Y. (2013). The vertical spatial metaphor of moral concepts and its influence on cognition. *ActaPsychologica Sinica, 45*(5), 538-545.

Wang, Z. H., Yao, Z. (2012). Concreteness effects of emotional noun words: Evidences from ERP. *ActaPsychologica Sinica, 44*(2), 154-165.

Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science, 322*(5901), 606-607.

Williams, L. E., Huang, J. Y., & Bargh, J. A. (2009). The scaffolded mind: Higher mental processes are grounded in early experience of the physical world. *European Journal of Social Psychology, 39*(7), 1257-1267.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*, 625-636.

Yan, S. C. (2011). Physical cleanliness and morality. *Advances in Psychological Science, 19*(8), 1242-1248.

Yang, J. P., Guo, X. M., &Wang, X. C. Metaphorical representation of moral concepts:Evidence from red/white color, left/right position and upright/skew font. *Acta Psychologica Sinica,49*(7),875-885.

Ye, H. S. (Ed). (2017). *Principles and applications of embodied cognition.* Beijing, China: The Commercial Press.

Yin, R., & Ye, H. S. (2014). The black and white metaphor representation of moral concepts and its influence on moral cognition. *Acta Psychologica Sinica, 46* (9), 1331-1346.

Yin, R., Su, D. Q., &Ye, H. S. (2013). Conceptual Metaphor Theory: Basing on Theories of Embodied Cognition. *Advances in Psychological Science, 21*(2), 220-234.

Yu, N., Wang, T, F., & He, Y. L. (2016). Spatial subsystem of moral metaphors: A cognitive semantic Study. *Metaphor and Symbol, 31*(4), 195-211.

Zhao, Y., Wu, L. (2019). The representation of moral metaphor in the vision of embodied cognition. *Psychological Exploration, 39*(4), 308-313.

Zhong, C. B., & Leonardelli, G. J. (2008). Cold and lonely: Does social exclusion literally feel cold? *Psychological Science, 19*(9), 838-842.

Zhu, Y., Zhang, L. (2001). Experimental research on the self-reference effect. *Science in China, 31*(6), 537-543.

Note: Figure translations are in progress. See original paper for figures.