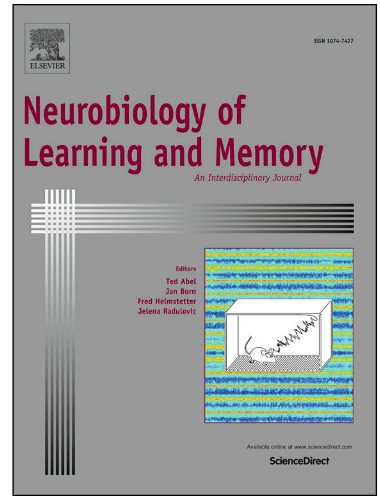
Journal Pre-proofs



Stimulus modality influences the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning

Jie Wu, Qiufang Fu, Michael Rose

PII: S1074-7427(19)30219-9
DOI: https://doi.org/10.1016/j.nlm.2019.107152
Reference: YNLME 107152

To appear in: *Neurobiology of Learning and Memory*

Received Date: 3 July 2019
Revised Date: 5 December 2019
Accepted Date: 23 December 2019

Please cite this article as: Wu, J., Fu, Q., Rose, M., Stimulus modality influences the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning, *Neurobiology of Learning and Memory* (2019), doi: https://doi.org/10.1016/j.nlm.2019.107152

Stimulus modality influences the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning

Jie Wu[1,2], Qiufang Fu[1,2], Michael Rose[3]

[1] State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[2] Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

[3] NeuroImage Nord, Department of Systems Neuroscience, University Medical Center Hamburg Eppendorf, Hamburg, Germany

Address for correspondence:

Qiufang Fu Ph. D

Institute of Psychology, Chinese Academy of Sciences

16 lincui Road, Chaoyang District, Beijing 100101, China

Tel: (86 10) 6484-5395

Fax: (86 10) 6487-2070

E-mail: fuqf@psych.ac.cn

Abstract

This study aimed to investigate whether stimulus modality influenced the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning and whether the use of the two strategies was supported by shared or separate neural substrates. To address these issues, we combined behavioral and fNIRS methods in a modified prototype distortion task in which each category member has one rule feature and ten similarity features, and each type of feature can be presented in either the visual modality or the auditory modality. The results in Experiment 1 revealed that the learning effect in the "auditory rule-visual similarity" condition was the highest among all four conditions; further analysis revealed that in the "auditory rule-visual similarity" condition, the number of participants who used the rule-based strategy was more than the number of participants who used the similarity-based strategy, and the learning effect was always much higher for the rule-based strategy than for the similarity-based strategy. The behavioral results in Experiment 2 replicated the main findings in Experiment 1, and the fNIRS results showed that the use of the visual rule-based strategy was mediated by the dorsolateral prefrontal cortex, whereas the use of the auditory similarity-based strategy mainly engaged in the superior temporal gyrus, and the use of the visual similarity-based strategy mainly engaged in the inferior temporal gyrus. The results in Experiment 3 revealed that when the stimuli had only one type of feature, the visual rule rather than the auditory rule was learned more easily. The results provide new evidence that the stimulus modality can influence the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning and that the use of the two types of strategies is supported by separate neural substrates both in the auditory modality and the visual modality.

Keywords: rule-based strategy, similarity-based strategy, stimulus modality, category learning

1. Introduction

Category learning is the process of establishing memory traces necessary to organize objects and events in the environment into separate classes (Ell, Ing, & Maddox, 2009). At least two types of strategies can be employed in category learning: the rule-based strategy, by which individuals categorize new objects primarily on the basis of rules, and the similarity-based strategy, by which individuals categorize new items mainly on the basis of their similarity to remembered exemplars or prototypes of the category (Smith & Minda, 1998; Smith., Patalano., & Jonides, 1998). The rule-based strategy was assumed to be formed explicitly through a declarative memory system (Carpenter, Wills, Benattayallah, & Milton, 2016; Koenig et al., 2005; Nomura et al., 2007) and to be employed when the category structure was defined by a verbal rule or prominent features (Ashby & Maddox, 2005; Ashby & Maddox, 2011; Carpenter et al., 2016; Koenig et al., 2005; Milton, Bealing, Carpenter, Bennattayallah, & Wills, 2017; Nomura et al., 2007). For example, if the square is red, it belongs to category A; if the square is blue, it belongs to category B. The similarity-based strategy was assumed to be formed implicitly through a non-declarative memory system (Koenig et al., 2005; Milton et al., 2017) and to be employed when the category structure was defined by the overall similarity to category examples or prototypes (Glass, Chotibut, Pacheco, Schnyer, & Maddox, 2012; Heindel, Festa, Ott, Landy, & Salmon, 2013; Koenig et al., 2005; Milton et al., 2017; Smith & Minda, 1998). For example, most objects within a category are small, round, and shiny, i.e., a family resemblance shared by those objects, and there is no single prominent feature that can be defined as a rule (Rabi, Miles, & Minda, 2015).

Controversy remains regarding whether the rule-based strategy and the similarity-based strategy were supported by multiple-category learning systems or a single system. Studies have suggested that the rule-based strategy and the similarity-based strategy are supported by different categories of learning systems with different neural substrates (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Maddox, 2005; Ashby & Maddox, 2011; Bozoki, Grossman, & Smith, 2006; Freedberg, Glass, Filoteo, Hazeltine, & Maddox, 2017; Koenig, Smith, Moore, Glosser, & Grossman, 2007; Nosofsky, Denton, Zaki, Murphy-Knudsen, & Unverzagt, 2012; Nosofsky & Zaki, 1998; Richler & Palmeri, 2014; Serre, 2016). For example, the use of the rule-based strategy has been demonstrated to be mediated by the anterior cingulate cortex, the thalamic and parietal region, and the left inferior prefrontal cortex (Koenig et al., 2005), whereas the use of the similarity-based strategy was mediated by the anterior prefrontal, posterior cingulate, and bilateral temporal-parietal regions (Koenig et al., 2005). Also demonstrated was that the rostro-lateral prefrontal cortex was important in abstracting and evaluating the rule in rule-based category learning (Paniukov & Davis, 2018), whereas the medial temporal lobe was essential in remembering and extracting examples in similarity-based category learning (Davis, Love, & Preston, 2012). However, some studies indicated that the use of the rule-based strategy and the similarity-based strategy might engage shared neural substrates (Carpenter et al., 2016; Edmunds, Milton, & Wills, 2018; Milton et al., 2017; Milton, Wills, & Hodgson, 2009; Newell, Dunn, & Kalish, 2011). For example, the processes of single-dimension sorting and overall similarity categorization was observed to activate the general cerebral regions, such as the left insula, the anterior cingulate cortex, and the left dorsolateral prefrontal cortex (Milton et al., 2009). The use of both the rule-based strategy and the similarity-based strategy was demonstrated to activate the prefrontal cortex and the occipital-temporal cortex (Milton et al., 2017).

To address this issue, researchers have examined whether certain factors could influence the acquisition and use of the two types of strategies in different manners. The acquisition and use of the two types of strategies could be modulated differently by practice (Thibaut, Gelaes, & Murphy, 2018; Verguts & Fias, 2009), age (Little & McDaniel, 2015; Rabi et al., 2015), culture (Murphy, Bosch, & Kim, 2017), and modality (Maddox, Ing, & Lauritzen, 2006). With practice, the process of category leaning could be indicated as a shift from abstraction-based learning (e.g., rule-based) to instance-based learning (Johansena & Palmeri, 2002) or from instance-based learning to abstraction-based learning (Johansena & Palmeri, 2002; Pothos, 2003; Raijmakers, Schmittmann, & Visser, 2014; Thibaut et al., 2018). For age, adults were observed to prefer the rule-based strategy, whereas children preferred the similarity-based strategy (Little & McDaniel, 2015; Murphy et al., 2017; Thibaut et al., 2018). For culture, European Americans and East Asians were observed to prefer the rule-based strategy and the similarity-based strategy, respectively (Gui et al., 2018; evidence for the opposite findings see Murphy et al., 2017). For modality, the stimulus modality was observed to influence the accuracy in the information integration category learning and the rule-based category learning, but in different manners (Yi & Chandrasekaran, 2016). For example, for the information-integration task, the learning effect was higher in the unimodal condition than in the multimodal condition (Maddox et al., 2006; Smith et al., 2014), but for the rule-based task, the learning effect was higher in the multimodal condition than in the unimodal condition, and the learning effect was superior in the auditory condition compared with the visual condition (Maddox et al., 2006). Nonetheless, whether the stimulus modality influences the acquisition and use of the rule-based strategy and the similarity-based strategy in multimodal and unimodal category learning remains unclear.

Therefore, this study aimed to investigate whether the stimulus modality could influence the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning, and whether the use of the rule-based strategy and the similarity-based strategy were supported by shared or separate neural substrates. To address the first issue, a modified prototype distortion task was adopted, in which each category member included one rule feature and ten similarity features, and each type of feature was presented in either the auditory modality or the visual modality. To address the second issue, the fNIRS technique was used, in which 56 channels were adopted to record the signals from the prefrontal cortex, the temporal lobe, and the visual cortex.

2. Experiment 1

The purpose of Experiment 1 was to explore whether the stimulus modality could influence the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning. To ensure that individuals could learn to classify stimuli by either the rule-based strategy or the similarity-based strategy, each stimulus had one rule feature and ten similarity features. To investigate the effect of the stimulus modality on the acquisition and use of the two strategies, each type of feature was presented in either the auditory modality or the visual modality. Thus, there were four conditions: the “auditory rule-visual similarity” (AR-VS) condition, in which the rule feature was presented in the auditory modality while the similarity features were presented in the visual modality; the “visual rule-auditory similarity” (VR-AS) condition, in which the rule feature was presented in the visual modality while the similarity features were presented in the auditory modality; the “visual rule-visual similarity” (VR-VS) condition, in which the rule feature and similarity features were presented both in the visual modality; and the “auditory rule-auditory similarity” (AR-AS) condition, in which the rule feature and similarity features were presented both in the auditory modality. To examine which type of the two strategies was dominantly acquired and used in category learning, the rule features from one category was recombined with the similarity features from the other category to form “ambiguous” items in the test phase. That is, the rule feature indicated membership in one category, and the similarity features indicated membership in the other category. Thus, according to the response proportions toward “ambiguous” items, we learned whether participants acquired and used the rule-based strategy or the similarity-based strategy in category learning.

2.1. Method

2.1.1. Participants

One hundred and fifty one undergraduate students (67 males and 84 females; mean age = 21.8 years, SD =2.29 years) voluntarily participated in this experiment. They were paid for their attendance. All of them had normal or corrected-to-normal vision. They were randomly assigned into four conditions. Data from two participants were excluded because their accuracy in the test phase was less than their respective group mean minus three standard deviations. There were 37, 38, 36, and 38 participants in the AR-VS, VR-AS, VR-VS, and AR-AS conditions, respectively. This experiment was approved by the committee for the protection of subjects at the Institute of Psychology, Chinese Academy of Sciences.

2.1.2. Materials

In the AR-VS condition, the auditory rule features comprised tones on piano including B, c, d, e, f, g, a, b, c^1,and d^1, i.e., low tones, and a^1, b^1, c^2, d^2, e^2, f^2, g^2, a^2, b^2, and c^3, i.e., high tones. Each tone was presented for 500 ms, at a rate of 44,100 HZ and 50 dB through dual-channel headphones. The visual similarity features were cartoon animals that varied on ten binary dimensions such as tail shape (feathery or pointy) and tail color (green or yellow), adopted from Gorlick and Maddox (2013). Each image was of 200*300 pixels, and the corresponding visual angle was 6.8° and 10.2°, respectively. Each category member had one rule feature and ten similarity features. During the training phase, there were ten members for each category (see Table 1). For example, the members of category A had one rule feature from L1to L10 (low tones), and seven "1" similarity features and three "0" similarity features of the ten binary dimensions (D1 to D10). The members of category B had one rule feature from H1to H10 (high tones) and seven "0" similarity features and three "1" similarity features of the ten binary dimensions (D1 to D10). During the test phase, there were 40 "ambiguous" items and 20 "training" items: the "ambiguous" items were created by recombining the rule and similarity features of stimuli from the two categories presented during the training phase, and the "training" items were created by recombining the rule and similarity features of stimuli from the same category presented during the training phase. For example, the "ambiguous" items had a combination of one rule feature from H1to H10, seven "1" similarity features, and three "0" similarity features, or had a combination of one rule feature from L1to L10, seven "0" similarity features, and three "1" similarity features. However, the "training" items had a combination of one rule feature from L1to L10, seven "1" similarity features, and three "0" similarity features, or had a combination of one rule feature from H1to H10, seven "0" similarity features, and three "1" similarity features. Thus, for both "ambiguous" and "training" items, the features were old but the combination was new. The only difference between the two types of items was that the rule and similarity features of the "ambiguous" items indicated different categories, and the rule and similarity features of the "training" items referred to the same category. The category structures of A and B were counterbalanced between participants.

Table 1. Category structure for each category member in the four conditions.

CA	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
Category A										
L1	1	1	0	0	1	0	1	1	1	1
L2	1	0	1	1	0	1	0	1	1	1
L3	0	1	0	1	1	1	1	0	1	1
L4	0	1	1	0	1	1	1	1	0	1

L5	1	0	1	1	0	1	1	1	1	0
L6	0	1	1	1	1	1	1	0	0	1
L7	1	0	1	1	1	1	0	1	1	0
L8	1	1	0	1	1	0	1	0	1	1
L9	1	1	1	0	1	0	1	1	0	1
L10	1	1	1	1	0	1	0	1	1	0
Category B										
H1	0	0	1	1	0	1	0	0	0	0
H2	0	1	0	0	1	0	1	0	0	0
H3	1	0	1	0	0	0	0	1	0	0
H4	1	0	0	1	0	0	0	0	1	0
H5	0	1	0	0	1	0	0	0	0	1
H6	1	0	0	0	0	0	0	1	1	0
H7	0	1	0	0	0	0	1	0	0	1
H8	0	0	1	0	0	1	0	1	0	0
H9	0	0	0	1	0	1	0	0	1	0
H10	0	0	0	0	1	0	1	0	0	1

Note: CA refers to rule features, and D1 to D10 refer to similarity features in the four conditions. L1 to L10 and H1to H10 are two sets of the visual rule features (low and high contrast) or the auditory-rule features (low and high pitch). D1to D10 is the visual or auditory similarity features. “0” and “1” refer to the specific feature of the binary dimension.

In the VR-AS condition, the visual rule feature comprised background images that varied in contrast from 13% to 40% (high contrast) and 58% to 85% (low contrast) with a 3% step. Each image was 330*350 pixels, and the corresponding visual angles were 11.2° and 11.9°, respectively. The auditory similarity features were musical melodies that varied on ten binary dimensions. Specifically, each melody comprised of five tones and each tone varied on timber (Chinese violin or guitar) or pitch (high or low). Each melody was presented for 2000 ms, in which each tone was presented for 400 ms. Auditory stimuli were presented at a rate of 44,100 HZ and 50 dB through dual-channel headphones. In the VR-VS and AR-AS conditions, the rule feature and similarity features were identical to the AR-VS and VR-AS conditions, except that both the rule feature and the similarity features of one stimulus were presented in the same modality.

2.1.3. Procedure

In the training phase, participants were asked to complete a classification task under one of the four conditions. As can be seen from Figure 1A, for each trial, a fixation was first presented for 500 ms. Next, the stimulus with one rule feature and ten similarity features was presented for 2000 ms, and participants were instructed to classify the cartoon animals or musical melodies as soon and as accurately as possible into category A or B depending on the conditions. Participants indicated their responses by pressing the keys “F” or “J”. The response keys were counterbalanced between participants. If the response was correct, correct feedback would be presented for 3000 ms; if the response was incorrect, incorrect feedback would be presented for 6000 ms. Twenty different stimuli (i.e., ten members of each category) were presented in a random sequence in each block. A short rest for at least 30 seconds was provided after each block. There were eight blocks, i.e., 20 different stimuli repeated eight times, for a total of 160 trials in the training phase.

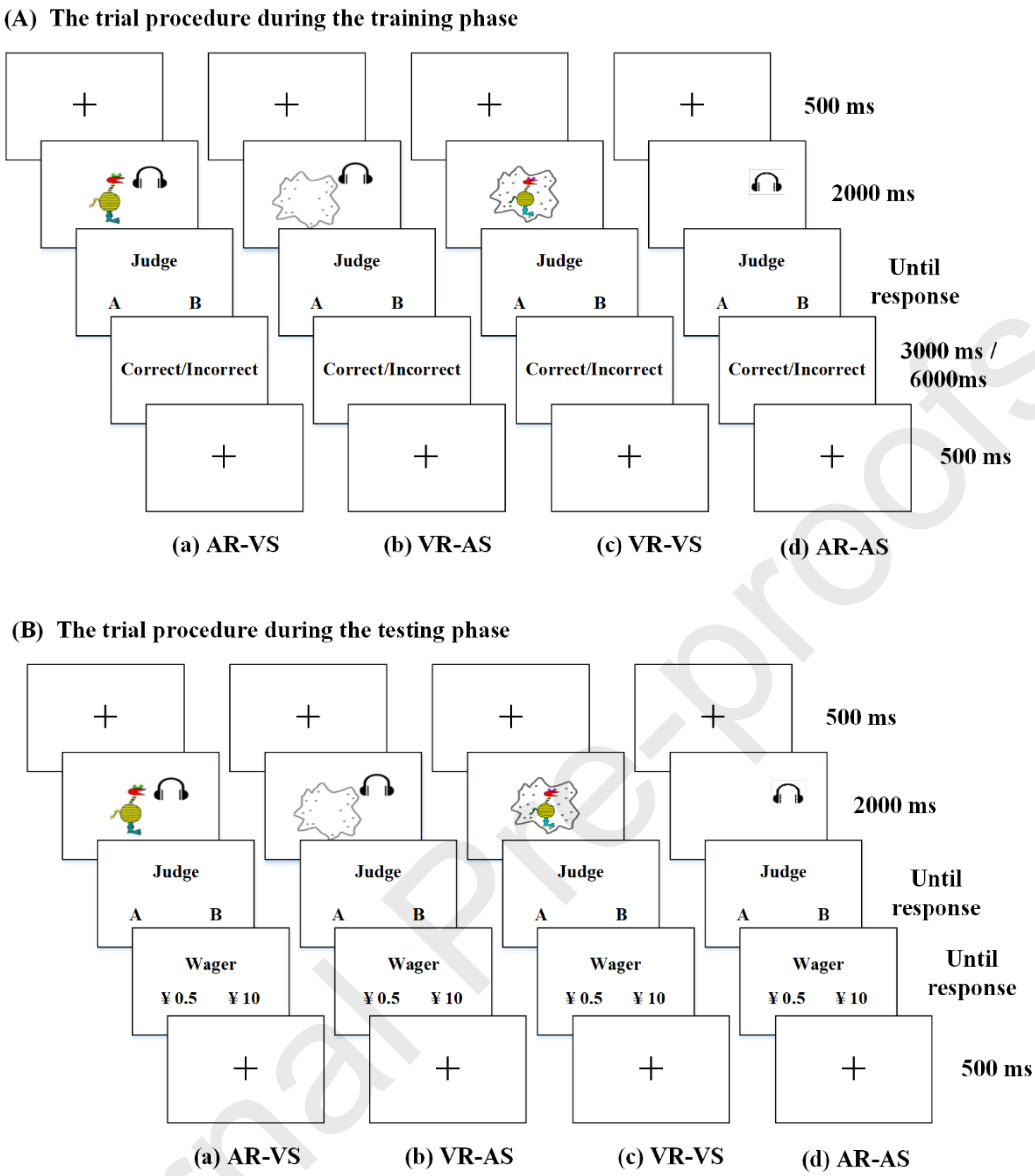


Figure 1. Trial procedures in Experiment 1. (A) Trial procedure during the training phase. (B) Trial procedure during the test phase.

In the test phase, participants were asked to complete a classification task on the basis of the knowledge they acquired in the training phase. As can be seen from Figure 1B, for each trial, a fixation was first presented for 500 ms. Next, the stimulus with one rule feature and ten similarity features was presented, and participants were asked to classify the cartoon animals or the musical melody into category A or B by pressing “F” or “J” depending on the condition. After their responses, participants were asked to make a bet of RMB 0.5 or RMB 10 based on their confidence in their response by pressing the key "F" and "J". Each block comprised 20 trials. The test phase comprised three blocks, in which 20 trials were "training" items and 40 trials were "ambiguous" items.

2.1.4. Results

2.1.4.1. Accuracy in the training phase.

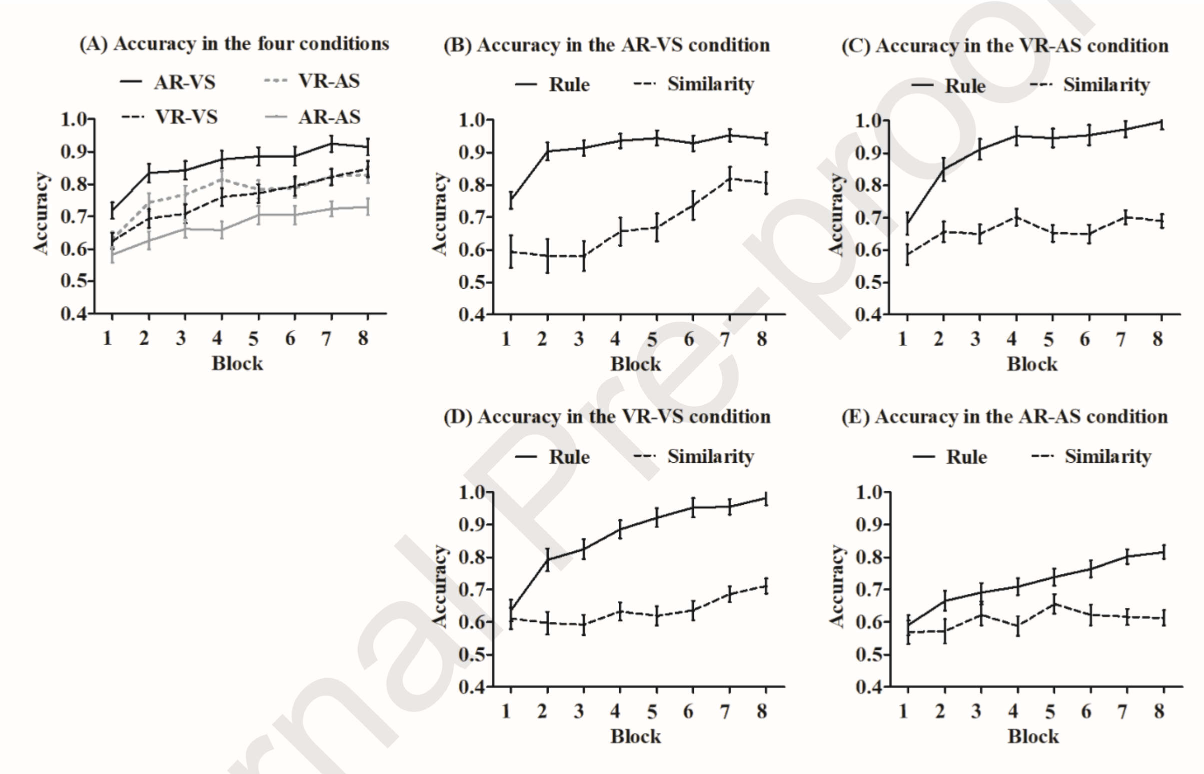


Figure 2. Accuracy in the training phase in Experiment 1. (A) Accuracy in the AR-VS, VR-AS, VR-VS, and AR-AS conditions. (B) Accuracy of the rule-based strategy group and the similarity-based strategy group in the AR-VS condition. (C) Accuracy of the rule-based strategy group and the similarity-based strategy group in the VR-AS condition. (D) Accuracy of the rule-based strategy group and the similarity-based strategy group in the VR-VS condition. (E) Accuracy of the rule-based strategy group and the similarity-based strategy group in the AR-AS condition.

Figure 2A shows the accuracy in the training phase in the four conditions. To examine whether individuals could acquire category knowledge in the four conditions, a 4 (stimulus modality: AR-VS vs. VR-AS vs. VR-VS vs. AR-AS) * 8 (blocks: 1 to 8) mixed ANOVA on accuracy was conducted. The result revealed that the main effect of blocks was significant, $F(7, 1015) = 49.10$, $p < .01$, $\eta_p^2 = .25$, indicating that the accuracy increased with training. The main effect of stimulus modality was also significant, $F(3, 145) = 13.19$, $p < .01$, $\eta_p^2 = .21$. The accuracy in the AR-VS condition was significantly higher than the other three conditions (*ps* < 0.01), and the accuracy in the AR-AS condition was significantly lower than the other three conditions (*ps* < 0.05). The interaction effect was not significant, $F(21, 1015) = 1.06$, $p = .40$, $\eta_p^2 = .02$.

According to how participants responded to the "ambiguous" items in the test phase, the participants in each condition could be further divided into two groups: one group mainly used the rule-based strategy, and the other group mainly used the similarity-based strategy. Figures 2B to 2E show the accuracy of the rule-based group and the similarity-based group in the four conditions, respectively. To explore how strategy influenced accuracy in the four conditions, a 4 (stimulus modality: AR-VS vs. VR-AS vs. VR-VS vs. AR-AS) * 2 (strategy group: rule-based vs. similarity-based) * 8 (blocks: 1 to 8) mixed ANOVA on accuracy was conducted. There was a significant effect of strategy group, $F(1, 141) = 221.34$, $p < .01$, $\eta_p^2 = .61$, indicating that the accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group. The interaction between the strategy group and the blocks was significant, $F(7, 987)= 9.11$, $p < .01$, $\eta_p^2 = .06$, and so was the three-way interaction, $F(21, 987)= 2.88$, $p < .01$, $\eta_p^2 = .06$. In the AR-VS and VR-AS conditions, the accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group in all blocks (*ps* < 0.05). In the VR-VS condition, the accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group from blocks two to eight (*ps* < 0.01). In the AR-AS condition, the accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group from blocks four to eight (*ps* < 0.05). The results suggested that the accuracy was much higher for the rule-based strategy than for the similarity-based strategy, and the learning advantage for the rule-based strategy occurred earlier in the multimodal conditions than in the unimodal conditions.

2.1.4.2. Accuracy in the test phase.

If a participant responded to "ambiguous" items in accordance with the rule-based strategy more often than the similarity-based strategy in the test phase, he or she was considered to have mainly used the rule-based strategy and was a member of the rule-based group and vise versa. Figure 3A shows the proportion of participants who mainly used the rule-based strategy and the similarity-based strategy in the four conditions. In the AR-VS condition, a chi-square test on proportion revealed that the number of participants (30 participants) who mainly used the rule-based strategy was significantly more than the number of participants (8 participants) who mainly used the similarity-based strategy, $X^2 = 12.74$, $df = 1$, $p < .01$, $\Phi = .59$. However, no significant differences were observed between the rule-based strategy group and the similarity-based strategy group in the other three conditions (VR-AS: 17 vs. 21 participants, $X^2 = 0.42$, $df = 1$, $p = .52$, $\Phi = .11$; VR-VS: 18 vs. 18 participants, $X^2 = 0.00$, $df = 1$, $p = 1$, $\Phi = 0$; VR-AS: 22 vs. 16 participants, $X^2 = 0.95$, $df = 1$, $p = .33$, $\Phi = .16$, respectively). The results revealed that more individuals preferred to use the rule-based strategy than the similarity-based strategy only in the AR-VS condition.

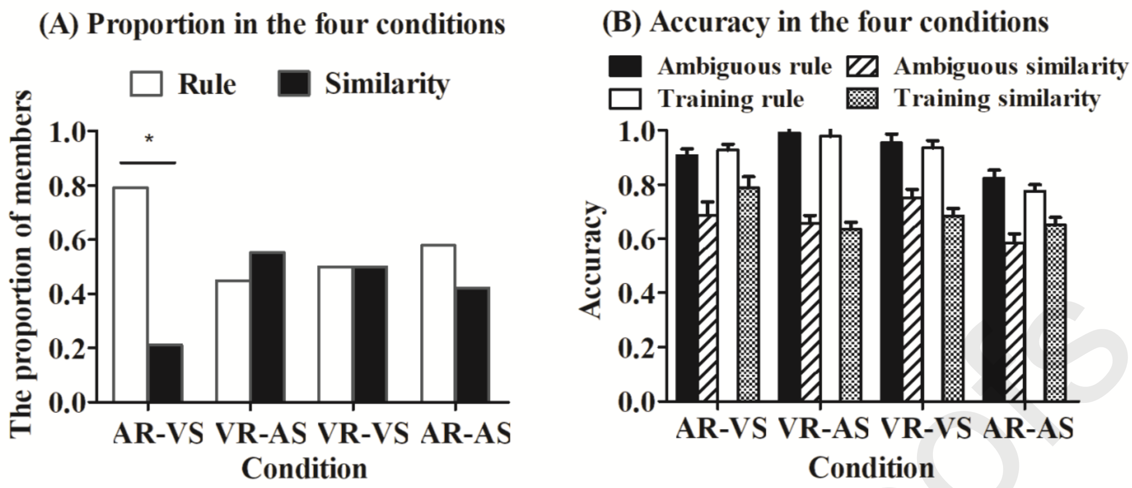


Figure 3. Proportion and accuracy in the test phase in Experiment 1. (A) Proportion of members in the four conditions. (B) Accuracy for the training and ambiguous items in the four conditions.

Figure 3B shows the accuracy for the "training" and "ambiguous" items in the four conditions. We calculated the accuracy for "ambiguous" items per participant based on their strategy group. That is, if a participant mainly used the rule-based strategy, the rule-based response would be correct, and vice versa. To explore how strategy influenced the accuracy for the training items in the test phase, a 4 (stimulus modality: AR-VS vs. VR-AS vs. VR-VS vs. AR-AS) * 2 (strategy group: rule-based vs. similarity-based) * 2 (items: ambiguous vs. training) mixed ANOVA was conducted. The results revealed a significant effect of the strategy group, F (1, 141) = 156.92, $p < .01$, $\eta_p^2 = .53$, indicating that the accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group. The main effect of modality was significant, F (3, 141) = 13.47, $p < .01$, $\eta_p^2 = .22$. Further analysis revealed that the accuracy in the AR-AS condition was significantly lower than the other three conditions ($p < .01$), but no significant difference was observed in any paired comparison among the AR-VS, VR-AS, and VR-VS conditions ($ps > .86$).

2.2. Discussion

The results showed that all participants could acquire the rule-based strategy or the similarity-based strategy in both multimodal and unimodal category learning, and the learning effects in the AR-VS condition were the highest among all four conditions. Further analysis revealed that more participants used the rule-based strategy compared with the similarity-based strategy in the AR-VS condition, but there was no strategy bias in the other three conditions. Notably, the accuracy of the rule-based strategy group was much higher than the similarity-based strategy group in all four conditions, and the learning advantage for the rule-based category learning occurred earlier in the multimodal conditions than in the unimodal conditions. The results indicated that the stimulus modality could influence the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning.

3. Experiment 2

The purpose of Experiment 2 was to further explore whether the stimulus modality could influence the acquisition and use of the rule-based strategy and the similarity-based strategy in category learning and whether the use of the two types of strategies was supported by shared or separate neural substrates. To address the first issue, Experiment 2 adopted a within-subjects design separately for the multimodal conditions and unimodal conditions. Specifically, participants in Experiment 2a completed the tasks in two multimodal conditions, i.e., the AR-VS and VR-AS conditions, and participants in Experiment 2b completed the tasks in two unimodal conditions, i.e., the VR-VS and AR-AS conditions. To address the second issue, we adopted fNIRS to examine the activation elicited by the two types of strategies in different cerebral cortexes. Studies have suggested that the use of the rule-based strategy was mediated by the prefrontal cortex (Koenig et al., 2005; Paniukov & Davis, 2018), while the use of the similarity-based strategy was supported by the lateral occipital region (Jiang et al., 2007). Thus, we placed 26 channels over the prefrontal cortex, nine channels over the temporal lobe, and 21 channels over the occipital cortex, respectively, for a total of 56 channels.

3.1. Experiment 2a

In Experiment 2a, participants completed the tasks in two multimodal conditions, i.e., the AR-VS and VR-AS conditions, in which the rule feature and the similarity features were presented in different modalities.

3.1.1. Method

3.1.1.1. Participants.

Forty-one young adult volunteers (19 males and 22 females; mean age = 21.7 years, SD = 2.76 years) voluntarily participated in this study. They were paid for their attendance. All of them had a normal or corrected-to-normal vision. This experiment was approved by the committee for the protection of subjects at the Institute of Psychology, Chinese Academy of Sciences.

3.1.1.2. Stimuli.

The stimuli of Experiment 2a were identical to the AR-VS and VR-AS conditions in Experiment 1, except that one block of "random" stimuli was added in the test phase. Each "random" stimulus included ten similarity features with five "0" features and five "1" features, and one rule feature including the central four tones of the central octave on the piano (e^1, f^1, $^{\#}f^1$, and g^1) in the AR-VS condition or the middle contrasts varied along 46%-55% with a 3% step in the VR-AS condition.

3.1.1.3. Procedure.

The trial procedures of Experiment 2a was identical to the AR-VS and VR-AS conditions in Experiment 1, except that after the response a blank was added for 1550 ms to

2450 ms, with a mean of 2000 ms. The training phase was shortened from eight blocks to five blocks. Participants were asked to complete the tasks in the two conditions sequentially. The sequence of the two conditions was counterbalanced between participants.

3.1.1.4 fNIRS data acquisition and data analysis

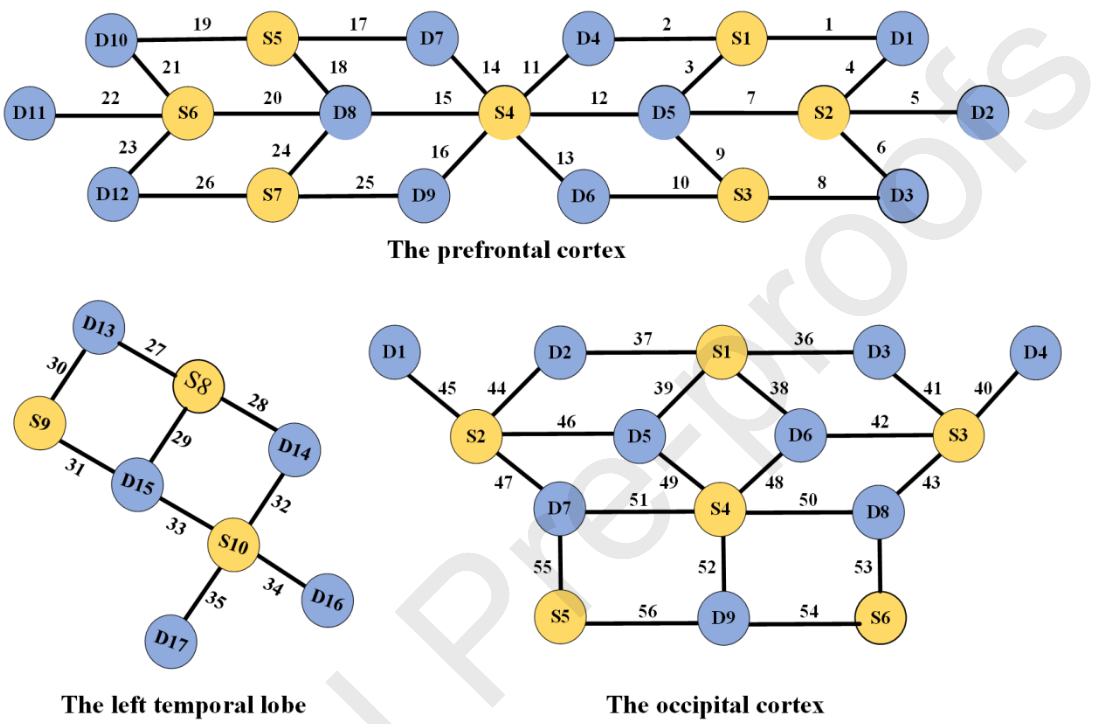


Figure 4. The channel layout over the prefrontal, temporal, and occipital cortexes

A multichannel commercial NIRS system was used to record Oxy-Hb and deoxy-Hb (NirsScan, Danyang Hui chuang Medical Equipment Co. Ltd). Measurements were performed on a continuous wave system using 16 sensors and 26 detectors, resulting in 56 channels placed over the prefrontal cortex, the temporal lobe, and the occipital cortex (Figure 4). According to the international 10-20 system, the S4 (the prefrontal cortex) was located over EEG coordinate FZ. S9 and D13 (the temporal lobe) were located over EEG coordinate FT7 and FC5, respectively. S4 (the occipital cortex) was located over EEG coordinate OZ. The channel distance was 30 mm, and the sampling rate was 10 Hz. The wavelengths were

780, 808, and 850 nm. According to the 3-D digital Talairach Altas, channels 1, 2, 3, 11, 14, 17, 18, and 19 were located over the frontal pole; channels 7, 8, 9, 10, 12,13, 15, 16, 20, 24, 25, and 26 were located over the dorsolateral prefrontal cortex; channels 4, 5, 6, 21, 22, and 23 were located over the pars triangularis Broca's area, channels 27, 28, and 32 were located over the superior temporal gyrus; channels 29, 30, and 31 were located over the middle temporal gyrus; channel 33 was located over the inferior temporal gyrus; channels 34 and 35 were located over the fusiform gyrus; channels 38, 39, 42, 43, 46, 48, 49, 50, and 51 were located over the V1; channels 36, 37, 41, 47, 52, 53, 54, 55, and 56 were located over the V2; and channels 40, 44, and 45 were located over the V3. The NIRS system has been used and verified in previous studies (Bu et al., 2018; Zhang et al., 2017).

The fNIRS data were exported and processed in the Nirs-SPM MATLAB toolbox. An event-related approach was used for data acquisition. Each trial lasted 9 s. According to the hemodynamic response function, each event was defined as the onset of the stimulus that lasts 4.5 s. The NIRS system recorded the Oxy-Hb and deoxy-Hb, which contributed to data analysis. According to the GLM modal, the coefficient of β was computed by comparing the observation modal with ideal modal. The NIRS data were mapped to the Montreal Neurological Institute (MNI) standard brain space. NIRS_SPM was used to define the region of interest based on Brodmann cytoarchitectonic areas coded in the 3-D digital Talairach Altas and the MNI template. Data from channels located over a corresponding brain region were averaged for subsequent analyses. To eliminate the overall noise, the wavelet-MDL was adopted; to eliminate the local time correlation, the data were digitally low-pass filtered with Gaussian. To illustrate the activation for using the rule-based strategy and the similarity-based strategy, a paired-samples t test was used to compare the activation averaged across channels for each brain region between the “training” items and the “random” items. To further illustrate whether the selected ROIs dominantly dealt with extracting and processing category knowledge or dealt with inconsistent information in decision making, a paired-samples *t* test was conducted to compare the activation across channels for each brain region between the training items and the ambiguous items.

3.1.2. Behavior results

3.1.2.1. Accuracy in the training phase

Figure 5A shows the accuracy in the AR-VS and VR-AS conditions. A 2 (stimulus modality: AR-VS vs. VR-AS) * 5 (blocks: 1 to 5) within-subjects ANOVA on accuracy was conducted. The result revealed only a significant main effect of blocks, $F(4, 160) = 27.34$, $p < .01$, $\eta_p^2 = .41$, indicating that accuracy increased with training. Neither the main effect of stimulus modality ($F(1, 40) = 0.059$, $p = .81$, $\eta_p^2 = .010$), nor the interaction effect was significant, ($F(4, 160) = 1.55$, $p = .19$, $\eta_p^2 = .037$).

As in Experiment 1, according to how participants responded to the "ambiguous" items, the participants in each condition were further divided into the rule-based group and the similarity-based group, Figures 5B and 5C show accuracy of the rule-based strategy group and the similarity-based strategy group in the AR-VS and VR-AS conditions, respectively. In the AR-VS condition, a 2 (strategy group: rule-based vs. similarity-based) * 5 (blocks: 1 to 5) mixed ANOVA revealed a significant effect of blocks, $F(4, 156) = 6.46$, $p < .01$, $\eta_p^2 = .14$. The main effect of the strategy group was also significant, $F(1, 39) = 24.82$, $p < .01$, $\eta_p^2 = .39$, indicating that accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group. The interaction was not significant, $F(4, 156) = 0.77$, $p = .55$, $\eta_p^2 = .02$. In the VR-AS condition, a comparable ANOVA revealed a significant effect of blocks, $F(4, 156) = 16.00$, $p < .01$, $\eta_p^2 = .29$, a significant strategy group effect, $F(1, 39) = 47.48$, $p < .01$, $\eta_p^2 = .55$, and a significant interaction, $F(4, 156) = 4.33$, $p < .01$, $\eta_p^2 = .10$. Further analysis revealed that accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group across all blocks (*ps* < 0.05).

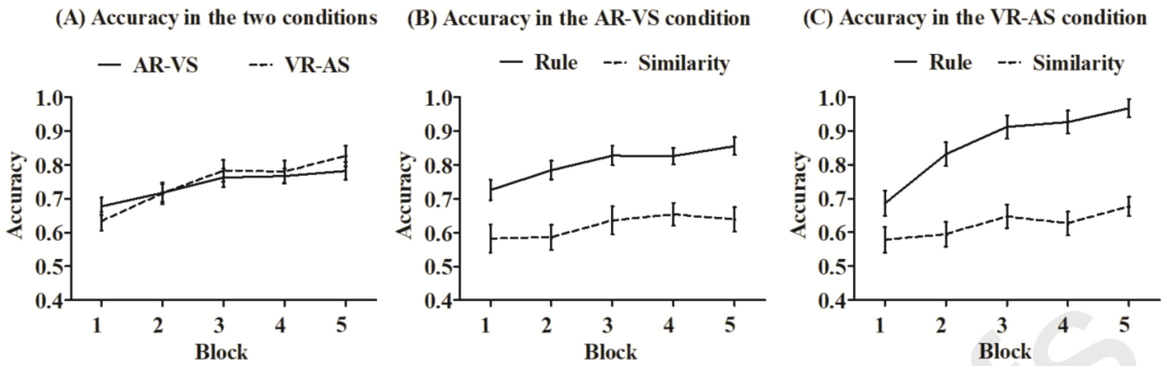


Figure 5. Accuracy in the training phase in Experiment 2a. (A) Accuracy in the AR-VS and VR-AS conditions. (B) Accuracy of the rule-based strategy group and the similarity-based strategy group in the AR-VS condition. (C) Accuracy of the rule-based strategy group and the similarity-based strategy group in the VR-AS condition.

3.1.2.2. Accuracy in the test phase

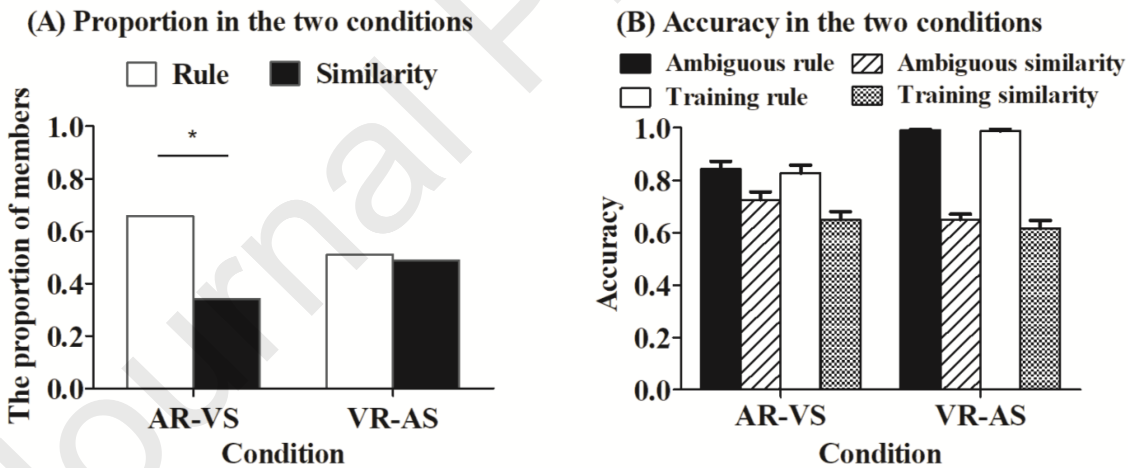


Figure 6. Proportion and accuracy in the test phase in Experiment 2a. (A) Response proportion for each strategy in the AR-VS and VR-AS conditions. (B) Accuracy for the "training" and "ambiguous" items in the two conditions.

Figure 6A shows the response proportion for each strategy group in the AR-VS and VR-AS conditions. As in Experiment 1, in the AR-VS condition, the chi-square test revealed that the number of members in the rule-based strategy group (27 participants) was significantly more than the number of members in the similarity-based strategy group (14 participants), $X^2 = 4.12$, $df = 1$, $p < .05$, $\Phi = 0.32$. However, no significant differences were observed between them (21 vs. 20 participants) in the VR-AS condition, $X^2 = 0.02$, $df = 1$, $p = 0.88$, $\Phi = .02$.

Figure 6B shows the accuracy for the "training" and "ambiguous" items in the test phase in the AR-VS and VR-AS conditions. We calculated accuracy for "ambiguous" items the same way as in Experiment 1. To explore how strategy influenced accuracy in the test phase in the AR-VS and VR-AS conditions, a 2 (strategy group: rule-based vs. similarity-based) * 2 (items: ambiguous vs. training) mixed ANOVA was used separately for each condition. In the AR-VS condition, it revealed a main effect of items, $F(1, 39) = 5.95$, $p < .05$, $\eta_p^2 = .13$, and a main effect of strategy group, $F(1, 39) = 11.79$, $p < .01$, $\eta_p^2 = .23$. The interaction was not significant, $F(1, 39) = 2.67$, $p = .11$, $\eta_p^2 = .06$. In the VR-AS condition, the main effect of items was not significant, $F(1, 39) = 1.56$, $p = .22$, $\eta_p^2 = .04$, but the main effect of strategy group was significant, $F(1, 39) = 372.94$, $p < .01$, $\eta_p^2 = .91$. The interaction was not significant, $F(1, 39) = 1.21$, $p = .28$, $\eta_p^2 = .03$. The results indicated that the accuracy was higher for the rule-based strategy group than for the similarity-based strategy group in both the AR-VS and VR-AS conditions.

3.1.3. fNIRS results

Participants had to randomly classify the "random" stimuli, while they could use the strategy acquired in the training phase to classify the training items in the test phase. To investigate the neural substrates supporting the use of the rule-based strategy and the similarity-based strategy in the test phase, we compared the brain activation between the “training” items and “random” items in the auditory-rule and visual-similarity groups in the AR-VS condition, and the visual-rule and auditory-similarity strategy groups in the VR-AS condition (see Table 2). A paired-samples *t* test on the “training” items and “random” items was conducted on the value of β averaged across channels for each brain region in the test phase. Compared to the “random” items, the “training” items activated the dorsolateral prefrontal cortex in the visual-rule group, $t(20) = 2.28$, $p = 0.03$, $d = 0.65$, activated the inferior temporal gyrus in the visual-similarity group, $t(13) = -2.61$, $p = 0.02$, $d = 1.13$, and activated superior temporal gyrus in the auditory-similarity group, $t(20) = -2.52$, $p = 0.02$, $d = 0.57$. No significant activation difference was found in the auditory-rule group.

Table 2. Significantly brain region activated by “training” items compared to the “random” items in the four groups in Experiment 2a.

Multisensory condition	Significantly activated brain region	*t*	*p*	*d*
Visual-Rule group	Dorsolateral prefrontal cortex	2.28	0.03	0.65
Visual-Similarity group	Inferior temporal gyrus	-2.61	0.02	1.13
Auditory-Rule group	None			
Auditory-Similarity group	Superior temporal gyrus	-2.52	0.02	0.57

Moreover, to investigate whether the brain activation between different strategy groups was comparable, an independent-samples *t* test was conducted on the value of β for “training” trials between the rule-based group and the similarity-based group for each significantly activated brain region in each condition. The results revealed no significant differences in the inferior temporal in the AR-VS condition, $t(37) = -.48$, $p = .64$, $d = .12$. Similarly, the result revealed neither significant differences in the dorsolateral prefrontal cortex, $t(39) = 1.00$, $p = .33$, $d = -.31$, nor significant differences in the superior temporal gyrus, $t(39) = .00$, $p = 1.00$, $d = 1.42$.

Finally, for the “ambiguous” items, their rule feature and similarity features indicated different categories, i.e., the two types of features were incongruent; however, for the “training” items, their rule feature and similarity features referred to the same category, i.e., the two types of features were congruent. To investigate the neural substrates for the congruency effect, we compared the activation between the “training” items and “ambiguous” items in the test phase. No significant activation differences were observed between “training” and “ambiguous” items in the visual-rule, visual-similarity, and auditory-rule groups. However, in the auditory-similarity group, compared to the “ambiguous” items, the “training” items activated the inferior temporal gyrus, $t(20) = 3.31$, $p < 0.01$, $d = 0.71$.

3.2. Experiment 2b

In Experiment 2b, participants completed the tasks in two unimodal conditions, i.e., the VR-VS and AR-AS conditions, in which the two types of features were both presented in the same modality.

3.2.1. Method

3.2.1.1. Participants

Forty young adult volunteers (18 males and 22 females; mean age = 21.3 years, SD = 2.71 years) participated in this experiment and were paid for their attendance. All participants had a normal or corrected-to-normal vision. This experiment was approved by the committee for the protection of subjects at the Institute of Psychology, Chinese Academy of Sciences.

3.2.1.2. Stimuli

The stimuli of Experiment 2b were identical to the VR-VS and AR-AS conditions in Experiment 1, except that one block of “random” stimuli was added in each condition and the training phase was shortened from eight blocks to five blocks. The “random” stimulus structure was the same as in Experiment 2a.

3.2.1.3. Procedure

The trial procedures of Experiment 2b was identical to the VR-VS and AR-AS conditions in Experiment 1, except that one blank was added for 1550 ms to 2450 ms, with a mean of 2000 ms, after the response. Participants were asked to complete the tasks in the two conditions, sequentially. The sequence of the two conditions was counterbalanced between participants.

3.2.1.4 fNIRS data acquisition and data analysis

The fNIRS data acquisition and data analysis were identical to those in Experiment 1.

3.2.2. Behavior results

3.2.2.1. Accuracy in the training phase

Figure 7A shows the accuracy in the training phase in the VR-VS and AR-AS conditions. A 2 (stimulus modality: VR-VS vs. AR-AS) * 5 (blocks: 1 to 5) within-subjects ANOVA on accuracy was conducted. The results revealed only a main effect of blocks, $F(4, 156) = 31.87$, $p < .01$, $\eta_p^2 = .45$, indicating a learning effect. Neither the main effect of stimulus modality ($F(1, 39) = 0.007$, $p = .94$, $\eta_p^2 = .00$), nor the interaction was significant, ($F(4, 156) = 1.45$, $p = .22$, $\eta_p^2 = .036$).

Using the same procedure as in Experiment 1, the participants in each condition were further divided into the rule-based strategy group and the similarity-based strategy group. Figures 7B and 7C show the accuracy of the rule-based strategy group and similarity-based strategy group in the VR-VS and AR-AS conditions, respectively. In the VR-VS condition, a 2 (strategy group: rule-based vs. similarity-based) * 5 (blocks: 1 to 5) mixed ANOVA revealed a significant effect of blocks, $F(4, 152) = 22.01$, $p < .01$, $\eta_p^2 = .37$, and a significant strategy group effect, $F(1, 38) = 54.42$, $p < .01$, $\eta_p^2 = .59$, indicating that the accuracy was significantly higher for the rule-based strategy group than for the similarity-based strategy group. The interaction was not significant, $F(4, 152) = 0.65$, $p = .63$, $\eta_p^2 = .02$. In the AR-AS condition, a comparable ANOVA also revealed a significant effect of blocks, $F(4, 152) = 9.87$, $p < .01$, $\eta_p^2 = .21$, and a significant strategy group effect , $F(1, 38) = 6.23$, $p < .05$, $\eta_p^2 = .14$, confirming the advantage for the rule-based strategy. The interaction was not significant, $F(4, 152) = 2.27$, $p = .07$, $\eta_p^2 = .06$.

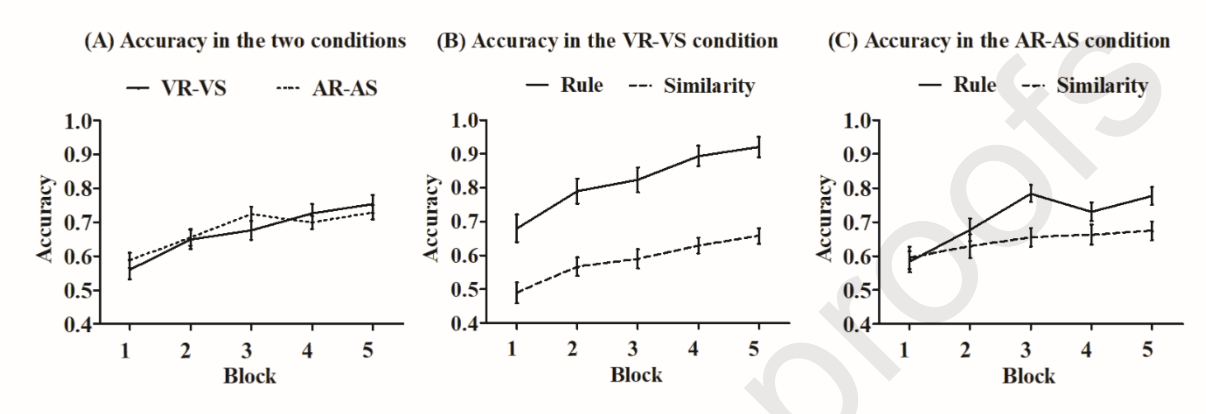


Figure 7. Accuracy in the training phase in Experiment 2b. (A) Accuracy in the AR-VS and VR-AS conditions. (B) Accuracy of the rule-based strategy group and the similarity-based strategy group in the VR-VS condition. (C) Accuracy of the rule-based strategy group and the similarity-based strategy group in the AR-AS condition.

3.2.2.2. Accuracy in the test phase

Figure 8A shows response proportions for each strategy in each unimodal condition. As in Experiment 1, the chi-square test revealed that there were no significant differences between the number of participants of the rule-based strategy group and the number of participants of the similarity-based strategy group in both the VR-VS condition (15 vs. 25 participants), $X^2 = 2.95$, $df = 1$, $p = .09$, $\Phi = .27$, and the AR-AS condition (21 vs. 19 participants), $X^2 = 0.22$, $df = 1$, $p = .64$, $\Phi = .07$.

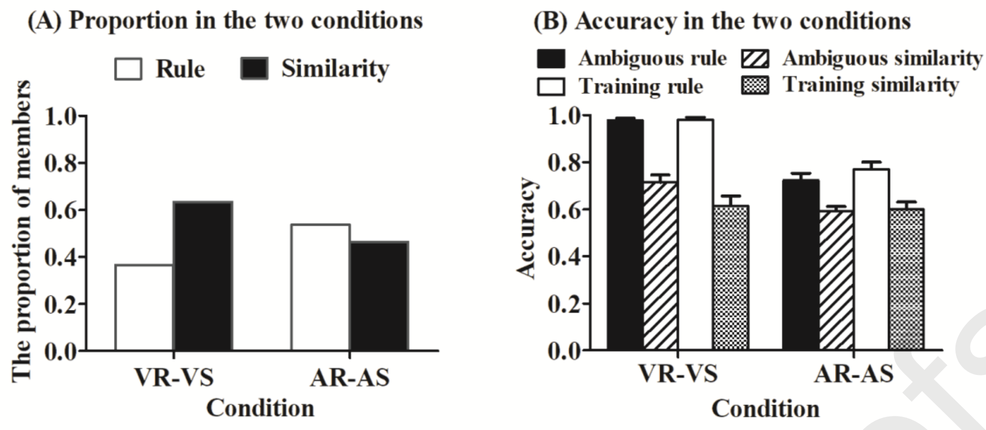


Figure 8. Proportion and accuracy in the test phase in Experiment 2b. (A) Response proportion of each strategy group in the VR-VS and AR-AS conditions. (B) Accuracy for the "training" and "ambiguous" items in the two conditions.

Figure 8B shows the accuracy for the "training" and "ambiguous" items in the test phase in the VR-VS and AR-AS conditions. We calculated accuracy for "ambiguous" items the same manners as in Experiment 1. To explore how strategy influenced the accuracy in the test phase, a 2 (strategy group: rule-based vs. similarity-based) * 2 (items: ambiguous vs. training) mixed ANOVA was conducted in each condition. In the VR-VS condition, it revealed a significant effect of strategy group, $F(1, 38) = 67.80$, $p < .01$, $\eta_p^2 = .64$. Neither the main effect of items, $F(1, 38) = 2.91$, $p = .10$, $\eta_p^2 = .07$, nor the interaction was significant, $F(1, 38) = 3.31$, $p = .08$, $\eta_p^2 = .08$. In the AR-AS condition, it revealed only a significant effect of strategy group, $F(1, 38) = 21.05$, $p < .01$, $\eta_p^2 = .36$. Neither the main effect of items, $F(1, 38) = 1.59$, $p = .22$, $\eta_p^2 = .04$, nor the interaction was significant, $F(1, 38) = 0.91$, $p = .35$, $\eta_p^2 = .02$.

3.2.3. fNIRS results.

As in Experiment 2a, we compared the "training" items with "random" items in the test phase in the visual-rule group and the visual-similarity group under the VR-VS condition, and the auditory rule group and the auditory similarity group under the AR-AS condition. Compared to the "random" items, the "training" items activated the dorsolateral prefrontal cortex in the visual-rule group, $t(14) = 2.42$, $p < 0.05$, $d = 0.69$, but no significant activation difference in the visual-similarity group in the VR-VS condition (see Table 3). Notably, no significant activation differences were observed between the "training" and "random" items in the AR-AS condition.

Table 3. Significantly brain region activated by "training" items compared to the "random" items in the four groups in Experiment 2b.

Multisensory condition	Significantly activated brain region	*t*	*p*	*d*
Visual-Rule group	Dorsolateral prefrontal cortex	2.42	0.03	0.69
Visual-Similarity group	None			
Auditory-Rule group	None			
Auditory-Similarity group	None			

Moreover, as in Experiment 2a, an independent sample *t* test was also conducted on the value of β for the "training" trials between the rule-based strategy group and the similarity-based strategy group within the dorsolateral prefrontal cortex. The results revealed that no significant difference, $t(38) = 1.17$, $p = 0.25$, $d = -0.38$.

Finally, we also compared the activation between the "training" items and the "ambiguous" items in the test phase under each condition. In the VR-VS condition, no significant activation difference was observed between the "training" items and the "ambiguous" items. In the AR-AS condition, compared to the "ambiguous" items, the "training" items activated the dorsolateral prefrontal cortex in the auditory-similarity group,

$t(18) = -2.30$, $p < 0.05$, $d = 0.85$, but no significant activation difference between them in the auditory-rule group was observed.

3.3. Discussion

The behavior results of Experiment 2 replicated the main findings in Experiment 1. Participants in the multimodal and unimodal conditions could learn the rule-based strategy or the similarity-based strategy in category learning, and the accuracy was much higher for the rule-based strategy group than for the similarity-based strategy group. Notably, more participants mainly used the rule-based strategy compared with the similarity-based strategy to complete the category task only in the AR-VS condition, confirming that the stimulus modality could influence the preference for a certain strategy.

The fNIRS results revealed that the use of the visual rule-based strategy mainly engaged the dorsolateral prefrontal cortex in the multimodal and unimodal conditions, while the use of the visual similarity-based strategy mainly engaged the inferior temporal gyrus and the use of the auditory similarity-based strategy mainly engaged the superior temporal gyrus. The results were consistent with studies that have suggested that the use of the visual rule-based strategy and the visual similarity-based strategy were supported by dissociative neural substrates (Koenig et al., 2005; Nomura et al., 2007). Although no significant activation differences were observed between the “training” and “random” items in the auditory rule-based strategy group in the multimodal and unimodal conditions, the use of the auditory similarity-based strategy engaged the superior temporal gyrus. The results were also consistent with a dual system, which supported the use of the two strategies in auditory categorization to some extent.

Significant activation differences were also observed in the dorsolateral prefrontal cortex between “training” items and “ambiguous” items in auditory similarity groups in the unimodal condition. Studies have suggested that the dorsolateral prefrontal cortex plays an critical role in rule-based categorization and similarity-based categorization (Carpenter et al., 2016; Edmunds et al., 2018; Milton et al., 2017; Milton et al., 2009; Newell et al., 2011). It might imply that although the dorsolateral prefrontal cortex was engaged in both rule-based categorization and similarity-based categorization, it played a different role in those two conditions. For rule-based categorization, the dorsolateral prefrontal cortex was related to processing and extracting rule-based knowledge, especially in the visual modality; for similarity-based categorization, it was involved in managing incongruent information in perception category knowledge, especially in the auditory modality.

4. Experiment 3

The results of Experiments 1 and 2 suggested that participants tended to use the rule-based strategy when the rule feature was presented in the auditory condition. Thus, an argument could be that this phenomenon occurred because the auditory rule was much easier to learn than the visual rule. To test this possibility, the stimuli presented in Experiment 3 included only one type of feature, and participants were asked to complete the visual rule condition (VR), the auditory rule condition (AR), the auditory similarity condition (AS), and the visual similarity condition (VS) separately. The sequence of the four conditions was counterbalanced between participants.

4.1. Method

4.1.1. Participants

Nineteen young adult volunteers (11 males and 8 females; mean age = 22.47 years, SD = 3.47 years) voluntarily participated in this experiment. All participants were paid for their attendance. All of them had a normal or corrected-to-normal vision. This experiment was approved by the committee for the protection of subjects at the Institute of Psychology, Chinese Academy of Sciences.

4.1.2. Materials and procedure

The stimuli were identical to Experiment 1, except that each stimulus included only one type of feature. That is, the rule-based feature was presented in the visual and auditory modalities, respectively, as were the similarity-based features.

The training phase comprised five blocks, and the test phase comprised one block, for a total of six blocks in each of the four conditions. Each block included 20 trials, including 120 trails in each of the four conditions.

4.2. Results

4.2.1. Accuracy in the training phase

Figure 9A shows the accuracy in the training phase in the four conditions. A 2 (stimulus modality: auditory vs. visual) * 2 (strategy: rule-based vs. similarity-based) * 5 (blocks: 1 to 5) within-subjects ANOVA revealed a main effect of stimulus modality, $F(1, 18) = 8.28$, $p < .01$, $\eta_p^2 = .32$, suggesting that the accuracy was significantly higher for the visual modality than for the auditory modality. The main effect of strategy was significant, $F(1, 18) = 454.63$, $p < .01$, $\eta_p^2 = .96$, indicating that accuracy was significantly higher for the rule-based strategy than for the similarity-based strategy. The main effect of the blocks was also significant, $F(4, 72) = 18.47$, $p < .01$, $\eta_p^2 = .51$, suggesting that accuracy gradually increased with practice. The interaction between stimulus modality and strategy was significant, $F(1, 18) = 40.13$, $p < .05$, $\eta_p^2 = .69$. Further analysis revealed that the accuracy of the rule-based strategy was significantly higher for the visual modality than for the auditory modality ($p < .01$) but was not for the similarity-based strategy ($p = .41$).

4.2.2. Accuracy in the test phase

Figure 9B shows the accuracy in the test phase in the four conditions. A 2 (stimulus modality: auditory vs. visual) * 2 (strategy: rule-based vs. similarity-based) within-subjects ANOVA revealed a significant effect of strategy, $F(1, 18) = 251.53$, $p < .01$, $\eta_p^2 = .93$, and a significant interaction, $F(1, 18) = 10.31$, $p < .01$, $\eta_p^2 = .36$. Further analysis revealed that the accuracy of the rule-based strategy group was significantly higher in the visual modality than in the auditory modality ($p < 0.01$), but not for the similarity-based strategy group ($p = 0.65$). The effect of modality was not significant, $F(1, 18) = 2.81$, $p = .11$, $\eta_p^2 = .14$.

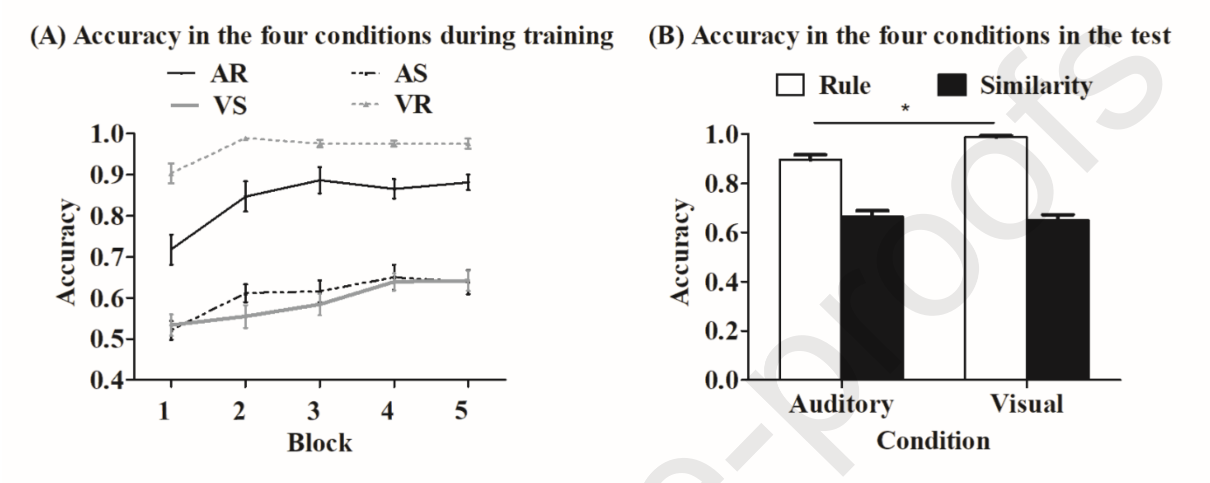


Figure 9. Accuracy in Experiment 3. (A) Accuracy in the training phase. (B) Accuracy in the test phase.

4.3. Discussion

As in Experiments 1 and 2, the results confirmed that the accuracy was significantly higher for the rule-based strategy than for the similarity-based strategy. However, the accuracy was significantly worse for the auditory rule-based task than for the visual rule-based task, and no significant difference was observed in accuracy between the auditory similarity-based task and visual similarity-based task. The results revealed that it was the auditory rule rather than the visual rule that was more difficult to learn when only the rule feature was presented. The results indicated that the preference for the auditory rule-based strategy in Experiments 1 and 2 was not because that the auditory rule was much easier to learn than the visual rule.

5. General discussion

The behavioral results indicated that individuals could learn the rule-based and similarity-based strategies in multimodal and unimodal category learning, and importantly, the stimulus modality could influence the acquisition and use of the rule-based strategy or the similarity-based strategy. Specifically, the behavioral results of Experiments 1 and 2 revealed that no bias in the rule-based strategy and the similarity-based strategy in the VR-AS, VR-VS, and AR-AS conditions, but more participants tended to use the auditory rule-based strategy than the visual similarity-based strategy to classify the stimuli in the AR-VS condition. Additionally, the accuracy was always better for the rule-based strategy than for the similarity-based strategy, and the learning advantage for the rule-based strategy occurred early in the multimodal conditions compared with the unimodal conditions. The behavioral results of Experiment 3 further indicated that the preference for the auditory rule-based strategy was not because the auditory rule was much easier to learn than the visual rule. Moreover, the fNIRS results indicated that the dorsolateral prefrontal cortex was involved in the visual rule-based categorization, the inferior temporal lobe was activated in the visual similarity-based categorization, and the superior temporal lobe was engaged in the auditory similarity-based categorization, providing further evidence that the use of the rule-based strategy and the similarity-based strategy was supported by separate neural substrates.

The behavioral results in the three experiments revealed that the accuracy gradually increased with training across all conditions, suggesting that individuals could learn the rule-based strategy and the similarity-based strategy no matter in what modality the rule feature and the similarity features were presented in the multimodal conditions and unimodal conditions. Studies have indicated that the rule-based strategy and the similarity-based strategy could be formed with practice in visual modality (Little & McDaniel, 2015; Murphy et al., 2017; Thibaut et al., 2018) and auditory modality (Alken, 1969; Maddox et al., 2014;

Maddox et al., 2006; Smith et al., 2014), respectively. Participants were observed to acquire visual or auditory category knowledge in the rule-based task and the information-integration task (Yi & Chandrasekaran, 2016). This study further extended this observation by demonstrating that the rule-based strategy and the similarity-based strategy could be acquired and used in multimodal and unimodal conditions.

According to how participants responded to the “ambiguous” items, participants in each condition were further divided into the rule-based group and the similarity-based group. The results in Experiments 1 and 2 indicated that accuracy was always much higher for the rule-based strategy group than for the similarity-based strategy group in the four conditions. To establish the rule-based strategy, participants must abstract the appropriate rule through hypothesis testing, which was resource-demanding. However, once participants obtained the correct rule, they would definitively know whether a stimulus was of a particular type or not and could perform very well. To establish the similarity-based strategy, participants must remember the category members or form a prototype. Because the comparison between the current stimulus and the remembered exemplars or prototypes is highly unstable, participants could have more difficult deciding which category the current stimulus belongs to (Kenett, Gold, & Faust, 2018). These findings were consistent in principle with studies that have suggested that the accuracy of the rule-based task was significantly higher than the accuracy of both the information-integration task (without a prominent and verbal rule) and the similarity-based task (Ell et al., 2009; Maddox et al., 2006; Wahlheim, McDaniel, & Little, 2016).

Notably, we found that the learning advantage for the rule-based category learning occurred earlier in the multimodal conditions than in the unimodal conditions and the number of participants who used the rule-based strategy was significantly more than the number of participants who used the similarity-based strategy in the AR-VS condition, but not in the

VR-AS, VR-VS and AR-AS conditions. Consistently, previous studies have found that for the rule-based task, the learning effect was higher in the multimodal conditions than in the unimodal conditions, and the accuracy was superior in the auditory modality than in the visual modality (Maddox et al., 2006). One might assume that this occurred because the auditory stimuli attracted more attention. If this was the case, the exception would be that there was a similar effect for the auditory similarity-based strategy. However, the results revealed that participants did not show any preference for the auditory similarity-based strategy, indicating that the preference for the auditory rule-based strategy might not be ascribed to more attention attracted by auditory stimuli. Additionally, the results in Experiment 3 showed the visual rule rather than the auditory rule could be learned more easily when the stimuli possessed only one feature, suggesting that the preference for the auditory rule might not be ascribed to the easiness of processing the auditory stimuli. Another possible explanation is that, for the visual objects, individuals tend to extract and encode more concrete and detailed information because of its abundant types of attributes such as shape, color, size, and etc.; while for the auditory objects, individuals tend to extract and encode more abstract and semantic information because of its obscurity and lack of variation. For example, daily objects often have the same auditory category label but different visual forms. Thus, the human brain might prefer the rule-based strategy for an auditory stimulus in the multimodal conditions.

The fNIRS results revealed that the dorsolateral prefrontal cortex was involved in the visual rule-based categorization, but not in the auditory rule-based categorization, and the visual similarity-based categorization and the auditory similarity-based categorization were mediated by the inferior temporal lobe and the superior temporal lobe, respectively. These findings were consistent with previous studies to some extent, which have indicated that the use of the rule-based strategy and the similarity-based strategy were support by separated category systems (Koenig et al., 2005; Maddox et al., 2013; Milton et al., 2017). The rule-based strategy required participants to form a prominent and verbal rule by hypothesis testing. Several studies have revealed that the prefrontal cortex was activated in the rule-based task (Carpenter et al., 2016; Koenig et al., 2005; Milton et al., 2017; Milton et al., 2009; Nomura et al., 2007) and the rostro-lateral prefrontal cortex played a crucial role in evaluating and switching the rule in the classification task (Paniukov & Davis, 2018). The similarity-based strategy required participants to compare the category members with a prototype or an exemplar. The temporal-parietal lobe was found to be activated in similarity-based categorization (Koenig et al., 2005; Milton et al., 2017; Milton et al., 2009).

However, our findings were inconsistent with the studies that have indicated that the rule-based strategy and the similarity-based strategy were supported by shared regions such as the dorsolateral prefrontal cortex (Carpenter et al., 2016; Edmunds et al., 2018; Milton et al., 2017; Milton et al., 2009; Newell et al., 2011). This might be because although the dorsolateral prefrontal cortex was involved in the rule-based strategy and the similarity-based strategy, it played a different role in the two strategies. For example, although no significant activation differences were observed in the dorsolateral prefrontal cortex between "training" items and "random" items for the auditory rule-based group, significant activation differences were observed between the "training" items and the "ambiguous" items for the similarity-based group, especially in the unimodal conditions. This result indicated that the dorsolateral prefrontal cortex might also be activated when inconsistent information needed to be included in decision-making.

6. Conclusion

This study found that individuals could learn the rule-based strategy and the similarity-based strategy in unimodal and multimodal conditions, and the learning effects were always much better for the rule-based strategy than for the similarity-based strategy. Notably, the learning advantage for the rule-based strategy occurred earlier in the multimodal conditions than in the unimodal conditions, and individuals preferred the rule-based strategy only in the AR-VS condition, suggesting that stimulus modality can influence the acquisition and use of the rule-based strategy and the similarity-based strategy. Furthermore, the fNIRS results revealed that the visual rule-based strategy was supported by the dorsolateral prefrontal cortex, and the visual similarity-based strategy was supported by the inferior temporal lobe. For the auditory condition, no significant activation difference was observed related to the auditory rule-based strategy, but the auditory similarity-based strategy was observed to involve the superior temporal lobe. These results illustrated that the use of the rule-based strategy and the similarity-based strategy are supported by separate neural substrates.

Acknowledgements

We are grateful to Danyang Hui chuang Medical Equipment Co. Ltd for their help with fNIRS data recordings and data analysis. The research was supported by the National Natural Science Foundation of China, the German Research Foundation (NSFC 61621136008/DFG TRR-169), and the National Natural Science Foundation of China (61632004).

References

Alken, E. G. (1969). Auditory discrimination learning Prototype storage and distinctive features detection mechanisms. *Perception & Psychophysics, 6*(2), 95-96. doi:10.3758/BF03210688.

Ashby, Alfonso-Reese, Turken, & Waldron. (1998). A Neuropsychological Theory of Multiple Systems in Category Learning. *Psychological Review, 105*(3), 442-481. doi:10.1037//0033-295X.105.3.442.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56*(1), 149-178. doi: 10.1146/annurev.psych.56.091103.070217.

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy Sciences, 1224*(1), 147-161. doi: 10.1111/j.1749-6632.2010.05874.

Bozoki, A., Grossman, M., & Smith, E. E. (2006). Can patients with Alzheimer's disease learn a category implicitly? *Neuropsychologia, 44*(5), 816-827. doi: 10.1016/j.neuropsychologia.2005.08.001.

Bu, L., Wang, D., Huo, C., Xu, G., Li, Z., & Li, J. (2018). Effects of poor sleep quality on brain functional connectivity revealed by wavelet-based coherence analysis using NIRS methods in elderly subjects. *Neuroscience Letters, 668*(6), 108-114. doi: 10.1016/j.neulet.2018.01.026.

Carpenter, K. L., Wills, A. J., Benattayallah, A., & Milton, F. (2016). A Comparison of the neural correlates that underlie rule-based and information-integration category learning. *Human Brain Mapping, 37*(10), 3557-3574. doi: 10.1002/hbm.23259.

Davis, T., Love, B. C., & Preston, A. R. (2012). Learning the exception to the rule: model-based FMRI reveals specialized representations for surprising category members. *Cerebral Cortex, 22*(2), 260-273. doi: 10.1093/cercor/bhr036.

Edmunds, C. E. R., Milton, F., & Wills, A. J. (2018). Due Process in Dual Process: Model-Recovery Simulations of Decision-Bound Strategy Analysis in Category Learning. *Cognitive Science, 42*, 833-860. doi: 10.1111/cogs.12607.

Ell, S. W., Ing, A. D., & Maddox, W. T. (2009). Criterial noise effects on rule-based category learning: the impact of delayed feedback. *Attention, Percepttion, & Psychophysics, 71*(6), 1263-1275. doi: 10.3758/APP.71.6.1263.

Freedberg, M., Glass, B., Filoteo, J. V., Hazeltine, E., & Maddox, W. T. (2017). Comparing the effects of positive and negative feedback in information-integration category learning. *Memory & Cognition, 45*(1), 12-25. doi: 10.3758/s13421-016-0638-3.

Glass, B. D., Chotibut, T., Pacheco, J., Schnyer, D. M., & Maddox, W. T. (2012). Normal aging and the dissociable prototype learning systems. *Psychology and Aging, 27*(1), 120-128. doi: 10.1037/a0024971.

Gorlick, M. A., & Maddox, W. T. (2013). Priming for performance: valence of emotional primes interact with dissociable prototype learning systems. *PLoS One, 8*(4), e60748. doi: 10.1371/journal.pone.0060748.

Gui, P., Li, J., Ku, Y., Li, L., Li, X., Zhou, X., . . . & Zhou, Y. D. (2018). Neural Correlates of Feedback Processing in Visuo-Tactile Crossmodal Paired-Associate Learning. *Frontiers in Human Neuroscience, 12*, 266. doi: 10.3389/fnhum.2018.00266.

Heindel, W. C., Festa, E. K., Ott, B. R., Landy, K. M., & Salmon, D. P. (2013). Prototype learning and dissociable categorization systems in Alzheimer's disease. *Neuropsychologia, 51*(9), 1699-1708. doi: 10.1016/j.neuropsychologia.2013.06.001.

Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., Vanmeter, J., & Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron, 53*(6), 891-903. doi: 10.1016/j.neuron.2007.02.015.

Johansena, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning. *Cognitive Psychology, 45*(2001),482-533. doi: 10.1016/S0010-0285(02)00505-4.

Kenett, Y. N., Gold, R., & Faust, M. (2018). Metaphor Comprehension in Low and High Creative Individuals. *Frontiers in Psychology, 9*, 482. doi: 10.3389/fpsyg.2018.00482.

Koenig, P., Smith, E. E., Glosser, G., DeVita, C., Moore, P., McMillan, C., . . . & Grossman, M. (2005). The neural basis for novel semantic categorization. *Neuroimage, 24*(2), 369-383. doi: 10.1016/j.neuroimage.2004.08.045.

Koenig, P., Smith, E. E., Moore, P., Glosser, G., & Grossman, M. (2007). Categorization of novel animals by patients with Alzheimer's disease and corticobasal degeneration. *Neuropsychology, 21*(2), 193-206. doi: 10.1037/0894-4105.21.2.193.

Little, J. L., & McDaniel, M. A. (2015). Individual differences in category learning: Memorization versus rule abstraction. *Memory & cognition, 43*(2), 283-297. doi: 10.3758/s13421-014-0475-1.

Maddox, W. T., Chandrasekaran, B., Smayda, K., & Yi, H. G. (2013). Dual systems of speech category learning across the lifespan. *Psychology and Aging, 28*(4), 1042-1056. doi: 10.1037/a0034969.

Maddox, W. T., Chandrasekaran, B., Smayda, K., Yi, H. G., Koslov, S., & Beevers, C. G. (2014). Elevated depressive symptoms enhance reflexive but not reflective auditory category learning. *Cortex, 58*, 186-198. doi: 10.1016/j.cortex.2014.06.013.

Maddox, W. T., Ing, A. D., & Lauritzen, J. S. (2006). Stimulus Modality Interacts with Category Structure in Perceptual Category Learning. *Perception & Psychophysics, 68*(7), 1176-1190. doi: 10.3758/BF03193719.

Milton, F., Bealing, P., Carpenter, K. L., Bennattayallah, A., & Wills, A. J. (2017). The Neural Correlates of Similarity- and Rule-based Generalization. *Journal of Cognitive Neuroscience, 29*(1), 150-166. doi: 10.1162/jocn_a_01024.

Milton, F., Wills, A. J., & Hodgson, T. L. (2009). The neural basis of overall similarity and single-dimension sorting. *Neuroimage, 46*(1), 319-326. doi: 10.1016/j.neuroimage.2009.01.043.

Murphy, G. L., Bosch, D. A., & Kim, S. (2017). Do Americans Have a Preference for Rule-Based Classification? *Cognitive Science, 41*(8), 2026-2052. doi: 10.1111/cogs.12463.

Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of Category Learning. *The Psychology of Learning and Motivation, 54*, 167-215. doi: 10.1016/b978-0-12-385527-5.00006-1.

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., . . . & Reber, P. J. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex, 17*(1), 37-43. doi: 10.1093/cercor/bhj122.

Nosofsky, R. M., Denton, S. E., Zaki, S. R., Murphy-Knudsen, A. F., & Unverzagt, F. W. (2012). Studies of implicit prototype extraction in patients with mild cognitive impairment and early Alzheimer's disease. *Journal of Experiment Psychology: Learning, Memory, and Cognition, 38*(4), 860-880. doi: 10.1037/a0028064.

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and nomal individuals: an exemplar-based interpretation. *Psychological Science, 9*(4), 247-255. doi: 10.1111/1467-9280.00051.

Paniukov, D., & Davis, T. (2018). The evaluative role of rostrolateral prefrontal cortex in rule-based category learning. *Neuroimage, 166*, 19-31. doi: 10.1016/j.neuroimage.2017.10.057.

Pothos, E. M. (2005). The Rules versus Similarity Distinction. *Behavioral and Brain Sciences, 28*(1),1-49.doi:10.1017/s0140525x05000014.

Rabi, R., Miles, S. J., & Minda, J. P. (2015). Learning categories via rules and similarity: comparing adults and children. *Journal of Experimental Child Psychology, 131*, 149-169. doi: 10.1016/j.jecp.2014.10.007.

Raijmakers, M. E., Schmittmann, V. D., & Visser, I. (2014). Costs and benefits of automatization in category learning of ill-defined rules. *Cognitive Psychology, 69*(1), 1-24. doi: 10.1016/j.cogpsych.2013.12.002.

Richler, J. J., & Palmeri, T. J. (2014). Visual category learning. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(1), 75-94. doi: 10.1002/wcs.1268.

Serre, T. (2016). Models of visual categorization. *Wiley Interdisciplinary Reviews: Cognitive Science, 7*(3), 197-213. doi: 10.1002/wcs.1385.

Smith, J. D., Johnston, J. J., Musgrave, R. D., Zakrzewski, A. C., Boomer, J., Church, B. A., & Ashby, F. G. (2014). Cross-modal information integration in category learning. *Attention, Perception, & Psychophysics, 76*(5), 1473-1484. doi: 10.3758/s13414-014-0659-6.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6),1411-1436. doi: 10.1037/0278-7393.24.6.1411.

Smith., E. E., Patalano., A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition, 65*(2-3), 167-196. doi:10.1016/S0010-0277(97)0043-7

Thibaut, J. P., Gelaes, S., & Murphy, G. L. (2018). Does practice in category learning increase rule use or exemplar use-or both? *Memory & Cognition, 46*(4), 530-543. doi: 10.3758/s13421-017-0782-4.

Verguts, T., & Fias, W. (2009). Similarity and rules United: similarity- and rule-based processing in a single neural network. *Cognitive Science, 33*(2), 243-259. doi: 10.1111/j.1551-6709.2009.01011.

Wahlheim, C. N., McDaniel, M. A., & Little, J. L. (2016). Category learning strategies in younger and older adults: Rule abstraction and memorization. *Psychology and Aging, 31*(4), 346-357. doi: 10.1037/pag0000083.

Yi, H. G., & Chandrasekaran, B. (2016). Auditory categories with separable decision boundaries are learned faster with full feedback than with minimal feedback. *The Journal of the Acoustical Socienty of America, 140*(2), 1332. doi: 10.1121/1.4961163.

Zhang, S., Zheng, Y., Wang, D., Wang, L., Ma, J., Zhang, J., . . . & Zhang, D. (2017). Application of a common spatial pattern-based algorithm for an fNIRS-based motor imagery brain-computer interface. *Neuroscience Letters, 655*(10), 35-40. doi: 10.1016/j.neulet.2017.06.044.

Highlights

- Learning is better for the rule-based strategy than the similarity-based strategy.
- Rule-based learning advantage occurs earlier in multi- than uni-modal conditions.
- Stimulus modality influences acquisition and use of the two types of strategies.
- The use of the two types of strategies is supported by separate neural substrates.