Spatial Categorization and Computation - Empirical Evidence from Artificial Label

Xianggang Xu
Institute of Psychology
Chinese Academy of Sciences,
Graduate University of Chinese Academy of Sciences
xuxg@psych.ac.cn

Xianghong Sun
Institute of Psychology
Chinese Academy of Sciences
sunxh@psych.ac.cn

Kan Zhang
Institute of Psychology
Chinese Academy of Sciences
zhangk@psych.ac.cn

Abstract

It is well known that spatial perception is a basic ability in our daily life, while we compute spatial relationship between two objects universally. This study examined how people perceive spatial categories using three tasks, learning task, producing task, and rating task. Three different kinds of spatial configurations were manipulated. 27 subjects were assigned randomly to each kind of spatial configuration. The results showed that response times(RTs) for spatial category judgment were, just as we have expected, slower near the boundary and faster near prototype, in learning task. RT pattern, in rating task, also suggested that people formed the expected configurations. People's rating patterns were well captured the spatial configurations predefined. Performances in all three tasks suggested that people can form artificial spatial categories and make spatial computation with enough practices. Implication for Bayesian analysis and application were briefly discussed.

Keywords: spatial perception, spatial categorization, Bayesian analysis.

1. Introduction

Concept learning and categorization have been studied for many decades. There are mainly four computational theories trying to model learning process, such as rule theory[1-3], prototype theory[4, 5], exemplar theory[6, 7], and boundary theory[8, 9]. Comparison and relationship between these modeling practices could be projected onto spatial domain. As we think, categorization in human is domain general, rather than domain specific. Therefore, concept learning and spatial perception may be grounded in common categorization mechanism from computational perspective[10, 11].

As for researches on spatial cognition, researchers found that people's estimation of spatial position was sensitive to spatial location where spatial projective language has high acceptability ratings [12]. Gapp (1994) also explored the basic meaning of spatial relation in 3D space[13]. Logan and Sadler (1996) found that people could project a spatial template to estimate the application of specific spatial terms [14]. Furthermore, Gapp (1995) found that people could select optimal reference to estimate object location [15].

Huttenlocher (1991) used a model of category effects (category adjustment model, CA) to interpret people's estimation of spatial location[16]. Their proposed estimation included truncation at category boundaries and weighting with a prototypical category value, while the stimuli own two levels of details: a fine-grain value and a category. This model were well justified by both stimulus judgment[17] and spatial reproduction task[18, 19]. Huttenlocher et al. (2004) found that people can not make use of distribution of the spatial locations in spatial reproduction. This result suggested that people use vertical and horizontal category boundaries to maximize the accuracy of estimates, and this spatial organization is very hard to change. Meanwhile, Spencer and Hund (2002) found that people use geometric cues to form spatial categories, and that adults' estimations are biased away from spatial boundary and toward geometric center. This result also suggested that common categorization processes may exist across spatial and object domains.

978-0-7695-3736-8/09 $25.00 © 2009 IEEE
DOI 10.1109/ICNC.2009.336

However, Simmering and Spencer (2007) found that subjects couldn’t use spatial category boundary without perceptual support[20].

Our study was to explore whether people could form the spatial categories they encountered, and the way how they form the spatial categories. In our study, subjects were asked to learn three different kinds of configurations. The perceived configurations were compared with the predefined ones after they completed three tasks in order. According to Huttenlocher’s CA model, the judgment of locations near the spatial boundary should entail much more uncertainty than judgment of those near the spatial prototype. Therefore, the RT of locations near the boundary should be slower than those near the prototype. Furthermore, acceptability near spatial prototype for a specific category should be better than those near spatial boundary.

2. Methods

2.1. Participants and design

Twenty-seven students participated in our study. All participants were recruited from universities in Beijing, China. All of them are with normal acuity, and right-handed. They were recompensed with 30RMB for one hour work. Three tasks were performed on a Lenovo desktop with 17 inch screen (with 1024×768 pixels), with which, subject seated before the screen comfortably. Participants were assigned randomly to one of three conditions.

Three different spatial configurations were manipulated, in which half circular area is enclosed with a semicircular and horizontal diameter. For each configuration, we used bound/bounds to divide the semi-circular area into two/three parts, named with artificial labels, dax, kav, and waf, resepctively. Both two-category configurations are only different at the position of bound, one for 90 degree radius, anther for 120 degree radius. The center for each category of two configurations is presented in table 1 and table 2. The unequal three-category configuration is presented in table 3. We named three kinds of configurations as equal two-category and unequal two-category, and unequal three-category, respectively. The sketches for these configurations were show in Figure 1.

Table 1. Equal two-category

spatial category	center	lower limit	upper limit
dax	135	120	150
kav	45	30	60
bound	90	75	105

Table 2. Unequal two-category

spatial category	center	lower limit	upper limit
dax	150	135	165
kav	60	45	75
bound	120	105	135

Table 3. Unequal three-category

spatial category	center	lower limit	upper limit
dax	150	138.75	161.25
kav	82.5	71.25	93.75
waf	22.5	11.25	33.75
bound1	45	33.75	56.25
bound2	120	108.75	131.25

There is one black spot in the center, and one red spot in the semi-circular area. The radius of circular area is 300 pixels, and its center point is on O(512, 434), whereas the red spot is sampled randomly for every possible position for each category on the screen. The coordinates of red spots will be recorded automatically by computer.

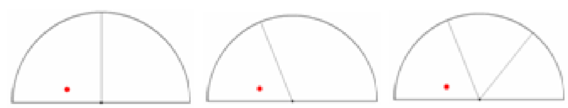

Figure 1. Three kinds of configurations. They were equal two-category and unequal two-category, and unequal three-category, from left to right, respectively. The spatial categorical labels were dax for the left category, and kav for the right, for both two-category configu-rations. The spatial categorical labels were dax, kav, and waf, from left to right, for unequal three-category.

2.2. Procedure

Each participant was asked to perform three tasks in turn, i.e. learning task, production task, and rating task. The first task, learning task, is a judgment task, in which subjects were asked to make a judgment that whether one randomized red spot belongs to an artificial categorical label (dax, kav, or waf). After that, a short period of some feedbacks will be shown to the subject on the screen. Then the subject was reminded to pay attention to the next trial. Feedbacks include the judgment correctness and speed of current trial and correct rate of all the foregoing trials they’ve done. All the responses and RTs will be recorded by the computer.

The second task, that is, production task, is very simple, in which subjects were asked to depict the best point of the corresponding categorical label with mouse. They were asked to move the mouse to their best expected position, and then click the left key of mouse. The position will be recorded by computer automatically.

For the last task, i.e. rating task, subjects were asked to give an acceptable rating of how category label match the corresponding spatial category of the red spot. The rating scale is a 9-point scale, ranging from 1 (not at all matching) to 9 (perfectly matching). The red spot was also randomized for all the trials, and all the ratings and rating times were recorded by computer.

3. Results

3.1. Learning task analysis

For the limited space, we only analyzed the RT of the last 30 accurate trials in the learning task. First of all, we coded those sampled points into different spatial types, boundary, prototype and bound(s). For both two-category configurations, we coded the category center and boundary area around the predefined values (15 degree), with corresponding lower limits and upper limits, which were shown in Table 1-2. Other points were coded as zero type. The similar coding method was applied to three-category configuration, only for 11.25 degree around predefined centers and bounds, as shown in Table 3.

For equal two-category configuration, although there is no significant difference among four types (F=1.79, p=0.15), bound type is a little slower (mean=1945.5, sd=754.71) than the other three types (mean_zero=1471.45, sd=894.69; mean_dax=1411.3, sd= 757.268; mean_kav=1554.19, sd=1383.11), which is shown in Figure 2.

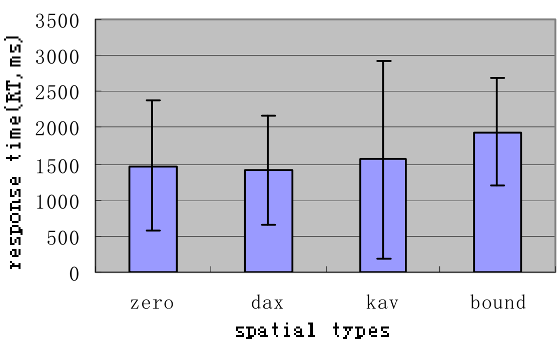

Figure 2. RT for each type in equal two-category

For unequal two-category configuration, as shown in Figure 3., significant difference (F=6.33, p<0.001) was found, in which, RT of bound type (mean_bound= 2162.83, sd=1242.29) was slower than that of the other three types (mean_zero=1499.07, sd=713.58; mean_dax= 1539.97, sd=861.08; mean_kav= 1432.84, sd = 577.89).

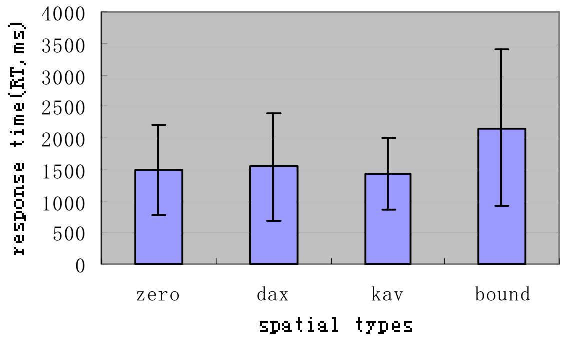

Figure 3. RT for each type in unequal two-category

For unequal three-category configuration, as shown in Figure 4., significant difference (F=6.15, p<0.001) was also found, that RT of two bound types was slower than that of the other four types. RT mean and standard deviation for each type were shown in Table 4.

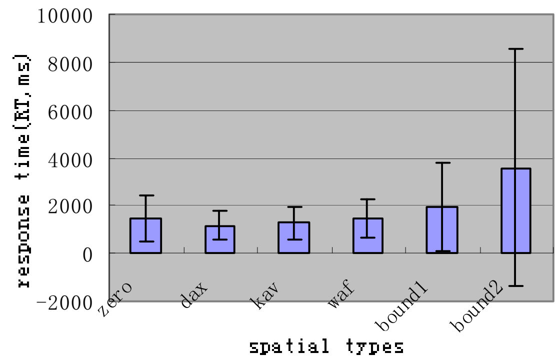

Figure 4. RT for each type in unequal three-category

Table 4. Means and standard deviations of RT for unequal three-category configuration.

spatial category	number	mean	std. deviation
zero	79	1436.95	977.42
dax	39	1165.33	599.83
kav	38	1286.58	691.32
waf	52	1475.58	823.70
bound1	20	1942.40	1831.81
bound2	12	3593.25	4984.43

3.2. Produce task analysis

3.2.1. Contours for spatial categorization

As subjects were asked to depict for just 10 times in equal two-category configuration, we only gave the contours for unequal two-category and unequal three-category configurations, as shown in Figure 5.

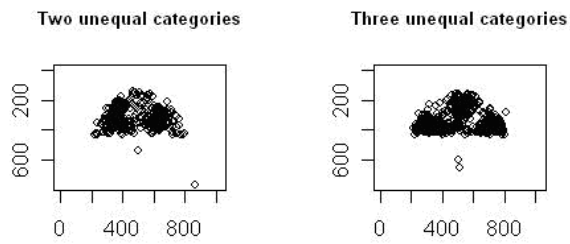

Figure 5. Depiction contours for two different spatial configurations, with unequal two-category for the left one and unequal three-category for the right one. The labels in X- and Y-axis were pixels.

The contours above show that people can form different spatial categories with enough learning procedure. Below, more details about comparison between perceived and predefined angle center for three different configurations were elaborated.

3.2.2 Comparison between perceived and actual configurations

We calculated the average degree of all the produced points for each category, and compared it with the predefined category center in the following table 5-7. With comparison between table 5 and table 6, the perceived angle center well captured predefined angle center of two categories configuration. Moreover, the perceived "kav" category almost perfectly matched the predefined angle center. Results in Table 7, compared with those in Table 5 and Table 6, show that subjects performed well in all three categories. These results imply that people formed different spatial configurations respectively.

Table 5. Comparison of perceived and predefined angle for equal two-category configuration

spatial category	perceived angle center	predefined angle center	deviance
dax	122.77	135	-12.23
kav	55.78	45	10.78

Table 6. Comparison of perceived and predefined angle for unequal two-category configuration

spatial category	perceived angle center	predefined angle center	deviance
dax	132.16	150	-17.84
kav	56.60	60	-3.40

Table 7. Comparison of perceived and predefined angle for three equal categories configuration

spatial category	perceived angle center	predefined angle center	deviance
dax	160.49	150	10.49
kav	91.06	82.5	8.56
waf	17.95	22.5	-4.55

3.3. Rate task analysis

In this part, we divided the trial into four stimulus types, dax-true, dax-false, kav-true, and kav-false, for two-category configurations, while dax-true, kav-true, waf-true, and all-false, for three-category configuration. The dependent variables are rating score and RT.

3.3.1 Rating score

The rating scores for equal two-category are significantly different between four types of trials (F=123.47, p<0.001), which is shown in Figure 6. Post hoc analysis shows that dax-true (mean=7.12, sd=2.32) and kav-true (mean=6.97, sd=2.32) types were rated higher than dax-false (mean=2.87, sd=2.44) and kav-false (mean=2.96, sd=2.31) types, while no differences between two true types and two false types.

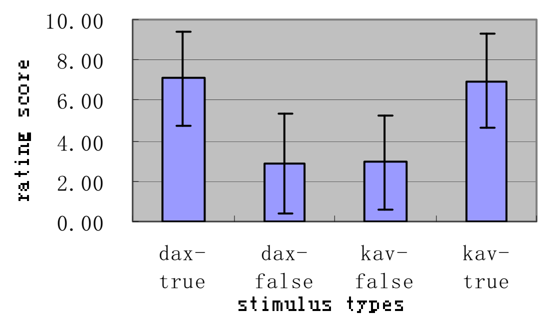

Figure 6. Rating scores for equal two-category.

Figure 7 shows the results of rating scores for unequal two-category. Similar to those in equal two-category, the rating scores were also significantly different between four types of trials (F=96.26, p<0.001). Post hoc analysis shows that dax-true (mean=6.98, sd= 2.81) and kav-true (mean=6.61, sd=23.02) types were rated higher than dax-false (mean=3.77, sd=3.16) and kav-false (mean=2.76, sd=3.01) types, while no differences between two true types and two false types.

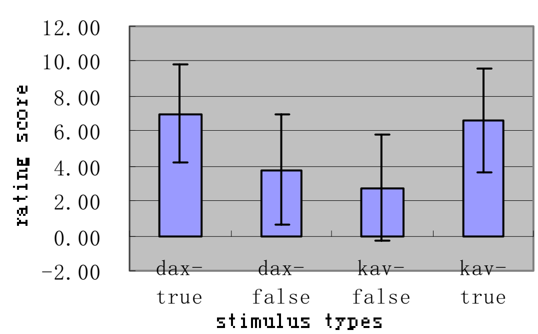

Figure 7. Rating scores for unequal two-category.

Figure 8 shows that all true types of trials were rated significantly different (F=823.78, p<0.001), with much higher for all true types than all-false trials (mean=2.04, sd=1.79). The means for dax-true, kav-true, and waf-true are 6.78, 7.19, 7.81, respectively, with corresponding sds, 2.26, 2.24, and 1.64.

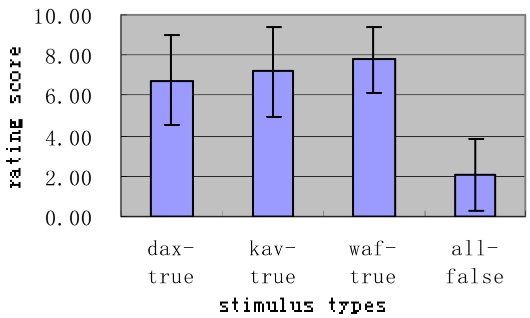

Figure 8. Rating scores for unequal three-category.

3.3.2 RT

RT analyses are shown in table 8-10 for equal two-category, unequal two-category, and unequal three-category, respectively. All three but the unequal two-category were not significantly different. As for unequal two-category, RT for dax-false was significantly longer than dax-true and kav-false, with ANOVA post hoc analysis (F=2.66, p=0.047).

Table 8. Mean and standard deviation for equal two-category.

type	RT mean	RT sd
dax-true	5601.49	5010.75
dax-false	7409.46	14052.73
kav-false	7570.49	7890.43
kav-true	6533.40	5813.31

Table 9. Mean and standard deviation for unequal two-category.

type	RT mean	RT sd
dax-true	3331.37	2295.64
dax-false	3900.83	2910.40
kav-false	3273.98	2094.36
kav-true	3433.66	2503.28

Table 10. Mean and standard deviation for unequal three-category.

type	RT mean	RT sd
dax-true	3669.88	2512.75
kav-true	3349.97	2420.67
waf-true	3174.94	2599.81
all-false	3365.10	2552.98

4. Discussions

Our study shows that people can form artificial spatial categories with enough learning practice. They performed differently as to spatial categorical boundaries and prototypes, in which RTs are fast near prototype than that near spatial boundary, as shown in the learning task. This finding implies that people can construct vague spatial ranges for certain objects implicitly, which also can be verified by production task performance in Figure 4. Using RT to examine spatial categorization performance is very promising, which we have never found in our domain of knowledge by now. To some extent, this finding extends the CA model for spatial category theory [16]. Moreover, with more encounters of category exemplars, we propose that people can impose a spatial boundary for spatial task. First, there is much geometric information, which can be integrated to update subject's belief of prototype and spatial category limits. Second, people's participation and judgment feedback encourage and facilitate them to calculate the spatial configuration actively.

As to rating task, people's rating pattern for three configurations unanimous confirms the formation of spatial categories. As evident from comparison of the four rating types, we can see that, if the label to be rated agrees with the spot position, then the rating score will be higher, and rating RT will be shorter. The reason for that pattern is that, contradictory rating trials will entail more cognitive resources, which increases the time for decision making. Basically, this analysis using four types for rating response is basically a signal detect paradigm.

Furthermore, we can get more information from individual learning process. For example, although people have no idea about the categorical labels at the beginning, they can give more and more correct judgments to situations where they encountered. This progress may be modeled through Bayesian approach[21, 22]. Machine learning results show that people can use some algorithms to implement the specific task. As for limited space, we may not discuss in detail here.

5. Acknowledgement

Work on this study was supported in part by the National Natural Science Foundation of China (Grant No. 30670701). We also thank two anonymous reviewers for their helpful comments and suggestions.

6. References

[1] Erickson, M. A. and Kruschke, J. K. Rule-based extrapolation in perceptual categorization. Psychonomic Bulletin & Review, 2002. 9(1): p. 160-168.

[2] Feldman, J., Minimization of Boolean complexity in human concept learning. Nature, 2000. 407(5): p. 630-633.
[3] Reed, S. K., Pattern recognition and categorization. Cognitive Psychology, 1972. 3(3): p. 382-407.
[4] Rosch, E., Cognitive Representations of Semantic Categories. Journal ol Experimental Psychology: General, 1975. 104(3): p. 192-233.
[5] Smith, J. D. and Minda, J. P. Journey to the Center of the Category: The Dissociation in Amnesia Between Categorization and Recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 2001. 27(4): p. 984-1002.
[6] Medin, D. L. and Schaffer, M. M. Context Theory of Classification Learning. Psychological Review, 1978. 85(3): p. 207-238.
[7] Nosofsky, R. M., Attention, Similarity, and the Identification-Categorization Relationship. Journal of Experimmatal Psychology: General, 1996. 115(1): p. 39-57.
[8] Ashby, F.G. and Gott, R. E. Decision Rules in the Perception and Categorization of Multidimensional Stimuli. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1988. 14(1): p. 33-53.
[9] Ashby, F. G., A Stochastic Version of General Recognition Theory. Journal of Mathematical Psychology, 2000. 44(2): p. 310-329.
[10]Shepard, R. N., Toward a Universal Law of Generalization for Psychological Science. Science, 1987. 237(4820): p. 1317-1323.
[11]Marr, D., Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. 1982: W. H. Freeman and Company.
[12]Hayward, W. G. and Tarr, M. J. Spatial language and spatial representation. Cognition, 1995. 55: p. 39-84.
[13]Gapp, K. P., Basic Meanings of Spatial Relations: Computation and Evaluation in 3D Space. Tech Report, 1994.
[14]Logan, G. D. and Sadler, D. D. A computational analysis of the apprehension of spatial relations, M.A.P.L.N. P.Bloom and M. F. Garrett, Editors. 1996, Cambridge, MA: MIT Press. p. 493-529.
[15]Gapp, K. P., Object Localization: Selection of Optimal Reference Objects. Tech report, 1995.
[16]Huttenlocher, J., Hedges, L. V. and Duncan, S. Categories and Particulars: Prototype Effects in Estimating Spatial Location. Psychological Review, 1991. 98(3): p. 352-376.
[17]Huttenlocher, J., Hedges, L. V. and Vevea, J. L. Why Do Categories Affect Stimulus Judgment? Journal of Experimental Psychology: General, 2000. 129(2): p. 220-241.
[18]Huttenlocher, J., et al., Spatial categories and the estimation of location. Cognition, 2004. 93: p. 75-97.
[19]Huttenlocher, J., et al., Estimating Stimuli From Contrasting Categories: Truncation Due to Boundaries. Journal of Experimental Psychology: General, 2007. 136(3): p. 502-519.
[20]Simmering, V. R. and Spencer, J. P. Carving Up Space at Imaginary Joints: Can People Mentally Impose Arbitrary Spatial Category Boundaries? Journal of Experimental Psychology: Human Perception and Performance, 2007. 33(4): p. 871-894.
[21]Feldman, J. and Singh, M. Bayesian estimation of the shape skeleton. PNAS, 2006. 103(47): p. 18014-18019.
[22]Xu, F. and Tenenbaum, J. B. Word Learning as Bayesian Inference. Psychological Review, 2007. 114(2): p. 245-272.