



# The Role of Top-Down Knowledge in Spatial Cueing Using Hierarchical Diagrams

Takeshi Sugio<sup>(✉)</sup>

Doshisha University, 1-3 Tatara-Miyakodani, Kyotanabe, Kyoto, Japan  
tsugio@mail.doshisha.ac.jp

**Abstract.** The way in which cueing a local element of a hierarchical diagram influences the distribution of visual attention was examined. Using a modified spatial cueing paradigm, the relation between the cue and the target was manipulated by two factors: the level at which the target was presented (higher, identical, lower), and the component to which the target belonged (same or different). The results showed an interaction between these two factors, and simple main effect analysis revealed that the detection time of the target was influenced by three factors: belongingness to the same component, geometrical collinearity of the nodes, and top bias, which regards the top of the diagram as being more informative. All of these factors are related to the conventional knowledge normally possessed about the particular category of the diagrams, and to how such knowledge affects the efficiency of the diagram comprehension process.

**Keywords:** Hierarchical diagram · Visual attention  
Conventional knowledge

## 1 Introduction

### 1.1 The Global Structure and Local Components of Hierarchical Diagrams

Schematic diagrams are essential tools for visual thinking, and hierarchical diagrams are representative examples of such diagrams [1]. They represent graphically the hierarchical relationship between items as a difference in levels, which are visually salient. In addition, they consist of the nodes that indicate the objects in the real world, and the line segments that connect each node to another at an either higher or lower level in the hierarchy. Some examples of hierarchical diagrams include: phylogenetic trees that show the evolutionary relationship between different plants and animals, tournament brackets that are often seen in sporting events, such as the Olympics, and organization charts that illustrate the relationships between the parts of and positions within an organization, such as a company. As seen in such examples, the levels in hierarchical diagrams may

represent abstract relations in both the spatial and temporal dimensions. Due to the visuospatial nature of such diagrams, we mistakenly feel that we can understand them intuitively without any prior knowledge in comparison with verbal description of the same data.

It has often been reported that visual attention is first directed to the global properties of stimuli, a phenomenon referred to as global precedence [2]. Therefore, when first glancing at a diagram, it is likely that we extract its global properties. The globality of the diagram's stimulus properties can be defined in terms of its level in the hierarchy of the data. In general, diagrams are more abstract than pictures [3]. Although they need not resemble the objects that they depict, the spatial and temporal relationships among the objects should be mirrored in the visuospatial structure of diagrams. Such relational information can be derived from multiple local elements of diagrams, so it occupies a higher level in the information hierarchy. Owing to the nature of diagrammatic representation, we can use relational information within the diagram to predict the possible behavior outcomes of the objects.

What kind of relational information in the diagrams, and how it is represented, differs according to the categories to which they belong [1]. For example, a matrix-type table represents combinatorial information between two different sets of items in an exhaustive manner, which is to say that the global structure of matrices is invariant with respect to changes in the actual existence of respective relations. If the problem that needs to be solved demands a close examination of the possible combinations of two sets of items, depicting the situation as a matrix is useful. Meanwhile, to obtain some information from the matrix, we need to have conventional knowledge of how it is constructed. By using such knowledge, we can guide our attention to the task-relevant relation effectively.

For hierarchical diagrams, their global shapes can be characterized in terms of the top and bottom sides of a circumscribed rectangle. Since the number of items decreases as the level in the hierarchy becomes higher, it is normally considered that a shorter side depicts a higher level. This property is characteristic of hierarchical diagrams in comparison with other types of schematic diagrams (i.e., matrices and networks, [1]). Whether such a property may act as a retrieval cue for the category to which the diagram belongs depends on the conventional knowledge of the observer.

On the other hand, the local components of the hierarchical diagrams need to be defined with care. The main purpose of reading a hierarchical diagram is not only to know what items are represented in the diagram, but also to obtain information about the hierarchical relationship among them. Therefore, neither a single node nor a line segment connecting multiple nodes may function as a component of a hierarchical diagram. According to Novick and Hurley, a building block of the hierarchical diagram consists of at least three nodes and two directional links connecting these nodes [1]. Such a block indicates the minimal unit of relational information. Thus, to obtain information from a hierarchical diagram in an effective way, the application of conventional knowledge is inevitable. Although appropriate knowledge can be activated in various ways, the global structure of the diagram serves as an effective cue for retrieval.

## 1.2 Effects of the Conventional Knowledge on Spatial Attention to Hierarchical Diagrams

As we have seen, conventional knowledge of a particular type of diagram plays an essential role in both its construction and comprehension. Empirical evidence showing the interaction between reading tasks and graph formats is compatible with this view [4]. The interaction between the bottom-up and top-down processes facilitates robust and efficient diagram comprehension. Trapp and Bar reviewed empirical findings showing the competitive nature of the perceptual process, and proposed a hypothesis that expectations were derived from the low spatial frequency component of visual images [5]. In the case of a diagram, conventional knowledge about how the information is visually organized may guide the attention of the viewer in an appropriate way. At the same time, perceptual cues such as the diagram's global shape or salient features may help activate the category to which the diagram belongs. In other words, both the bottom-up perceptual process and the use of top-down, conventional knowledge are involved in diagram comprehension, and understanding the underlying mechanisms of this comprehension may account for the effectiveness of diagrams in many situations.

Thus, when reading a diagram, visual attention is guided by both salient features and top-down knowledge. Understanding the attentional process in reading a diagram is important in that it shows both temporal and spatial range of information processing capabilities involved. There is also an empirical evidence that spatial attention is required for semantic processing [6]. Furthermore, studying the attentional mechanism of diagram reading could enable us to obtain a better understanding of how the diagram should be designed. However, few experimental results about such a mechanism are available to date. In the present experiment, a modified version of the spatial cueing task was used to examine the conditions in which the benefits for cueing arise [7].

## 2 Experiment

The purpose of the present experiment was to examine how cueing a local element of a hierarchical diagram affects the distribution of visual attention in the diagram spatially. Abstract four-layer hierarchical diagrams were used as stimuli (Fig. 1). Participants were told to fixate on the center of the diagram, and their task was to detect changes in the luminance of the rectangle, which represented a particular item. Before the change in luminance, one of the rectangles was brightened to induce the orientation of visual attention. In valid trials, the change in luminance occurred at the cued rectangle. In invalid trials, the change in luminance took place at a non-cued rectangle. The differences between the cue and target, that is, the rectangle at which the luminance was changed, were manipulated with respect to the following two variables: target level in regard to the cue (higher, identical, lower), and belongingness to the same component (a V-shaped figure that consists of three nodes connected by directional lines). Since the distance between the cue and the target is averaged across conditions, any performance difference between the conditions should be attributed

to the differences in the objects that participants form internally with the stimuli. Within the same object, attention directed at some part of it automatically spreads, resulting in benefits in performance at positions other than the cued position.

## 2.1 Method

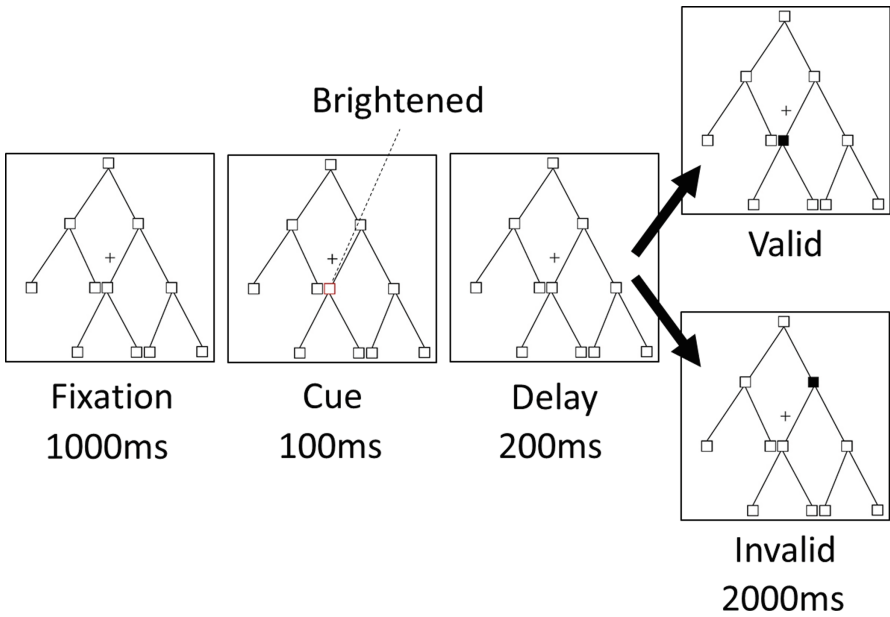
*Participants.* The participants were 20 undergraduate students at Doshisha University (five males and 15 females, mean age 21.9 years). All students were all paid for their participation, and all had normal or corrected vision.

*Apparatus.* Stimulus presentation and data collection were managed by SuperLab version 5 (Cedrus Corporation) running on a personal computer (HP Compaq Elite 8300SFF) with 17-inch cathode ray tube monitor (Iiyama, HF703U). Screen resolution was set to  $1280 \times 1024$  pixels during the experiment. The viewing distance was set as about 60 cm, and the participants' heads were stabilized using a chin rest. Responses were measured using a RB-530 response pad (Cedrus Corporation).

*Stimuli.* The stimuli were abstract hierarchical diagrams with rectangles as nodes (Fig. 1). All diagrams were composed of four layers, and each node was connected to two other nodes at a lower level. The fixation cross, stimuli, and target were gray, and the cue was white. The fixation cross was a plus sign, which subtended about  $0.04^\circ \times 0.04^\circ$ . Each stimulus subtended about  $10^\circ \times 10^\circ$  with a stroke of approximately  $0.02^\circ$ , and a square node subtended about  $0.04^\circ \times 0.04^\circ$ . The retinal distance between the cue and the target in the invalid condition was about  $2.5^\circ$ . The retinal distance between the fixation cross and the target was  $2.5^\circ$  on average, ranging from approximately  $1.0^\circ$  to  $5.0^\circ$ .

*Design.* The target appeared at either the same square as the cue (valid condition) or a different square from the cue (invalid condition). In the invalid condition, the target appeared at one of six different locations, that is the combinations of the level in the hierarchy compared to the cue (high, identical or low) and the belongingness to the same component as the cue (same or different). In total, 60% of the trials were valid, and 24% were invalid (4% for each condition). The remaining trials were catch trials in which no targets appeared.

*Procedure.* Each trial began with the presentation of overlapped stimuli (a fixation cross) for 1,000 ms, after which, the cue, brightening (i.e., a change in the luminance of the square from gray to white), was superimposed for 100 ms, and then returned to gray. The fixation display was presented again for 200 ms, and the target was presented (the square was filled-in). The target remained on the screen until a response was given, or, if there was no response, for 2,000 ms. Participants were told to respond by pressing the center button of a response pad as quickly as possible when they detected the target, and to withhold responses in the catch trials. A subsequent trial began after a 500-ms interval. The hierarchy



**Fig. 1.** The time course of a single trial in the experiment.

level at which the cue appeared was counterbalanced across trials to control the difference in density between above and below the cued level. The presentation sequence was randomized across participants. The time course of a single trial is shown in Fig. 1.

The participants were told that although their response latency would be recorded, it was important to minimize the number of errors. If a participant made an anticipatory response, defined as a response within 150 ms of the target presentation or a false alarm, a feedback beep was presented for 500 ms. Participants were also asked to maintain their focus on the fixation cross throughout each trial.

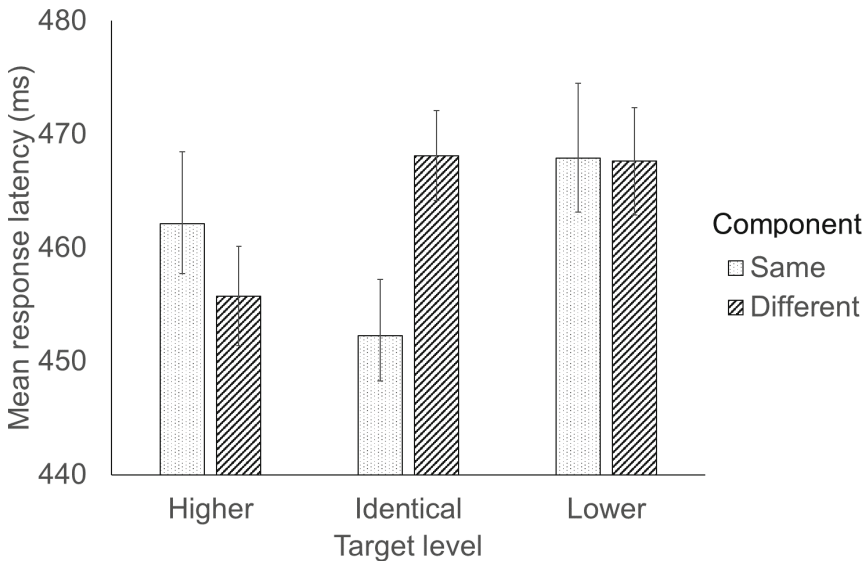
The experiment consisted of eight blocks of 125 trials each. Participants were allowed to take a rest between blocks if necessary. The task was explained at the beginning of each block and before the main trials, and 27 practice trials were given. If a participant could not respond correctly for 20 consecutive trials, a practice session was restarted, and the task was explained again. When there was no response to the target within 2,000 ms, when a response was made within 150 ms of the target response, or when the participant failed to withhold a response for the catch trial, the trial was considered as an error.

## 2.2 Results

The total error rate was 1.0%. The participants' median response latencies for correct trials under both the valid and invalid conditions were then compared

using a paired *t* test; the results showed a significant difference (valid condition = 411.350 ms vs. invalid condition = 461.325 ms;  $t(19) = 2.330, p = .031$ ).

The median response latencies for the invalid conditions were analyzed using repeated analysis of variance (ANOVA), with the within-participant factors of target level and component. The results are shown in Fig. 2. The main effect of target level and the interaction between two factors were significant (Greenhouse-Geisser adjusted results for target level:  $F(1, 19) = 4.055, p = .045, \eta_p^2 = .176$ ; interaction:  $F(2, 38) = 4.318, p = .020, \eta_p^2 = .185$ ). Multiple comparison for target level performed using Shaffer's method found that the lower condition took significantly longer than the higher or identical conditions (both  $p < .05$ ). As for the interaction, simple effect analysis revealed that the effect of level was marginally significant when both the cue and the target belonged to the same component (Greenhouse-Geisser adjusted  $F(2, 38) = 3.454, p = .058, \eta_p^2 = .154$ ), and highly significant when the target appeared at the node of a component different from the cue ( $F(2, 38) = 5.827, p = .006, \eta_p^2 = .235$ ). Multiple comparison for the simple effect under the same component condition showed a significant difference between the lower and identical levels ( $p < .05$ ), and that the higher-level condition was significantly faster than the lower- and identical-level conditions for the different component conditions (both  $p < .05$ ). The simple effect of the component was only significant when both the cue and the target were presented at the identical level in the hierarchical diagram ( $p < .05$ ).



**Fig. 2.** Interaction between the target level and component for response latencies.

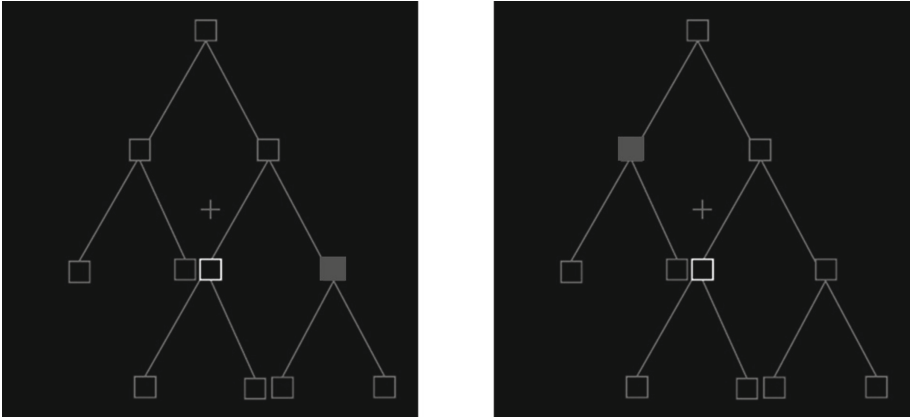
### 2.3 Discussion

The results clearly demonstrate that the effects of spatial cueing of a node in the hierarchical diagram differ according to the hierarchy level at which the target appears and the local V-shaped component to which the target belongs; both factors determine the informational structure of the hierarchical diagram. The present results support the claim that when viewing a hierarchical diagram, conventional knowledge about how the diagram is organized globally, and what constitutes its elements is activated automatically. Consequently, task-relevant information can be obtained from the diagram efficiently.

The main effect of target level showed that when a target appeared at a level lower than the cue, more time was needed to detect the change in luminance. This was not the case for the higher level, which suggests that we read the diagram from a level higher than the cued level in the hierarchy preferentially. Based on the global shape of the hierarchical diagram, we can easily determine which side is the top. We know that the top side of a hierarchical diagram depicts the items that indicate superordinate concepts, and such biases might help us comprehend the diagram.

The interaction between the target level and component is also important in that it suggests how nodes are organized as building blocks across different levels in the hierarchical diagram. When the target level was different from the cue, no significant difference was observed in detection time between the component conditions. On the other hand, when the target level was identical to the cue, the target was detected faster under the same component condition. This suggests that attending to a particular node in a diagram benefits from both the global structure and the local component of that diagram. Identifying the level in the hierarchy is essential in determining the representational range of the node, e.g., whether the node represents an animal or a fish. It may also be related to the entry level often discussed in the categorization literature [8]. By using the global features of a diagram, such as node collinearity, we may be able to identify the hierarchical structure and the level considered the entry point.

The examples in which the target was detected relatively faster are shown in Fig. 3. When the target was on the identical level as the cue and belonged to the same component (the left panel in Fig. 3), detection was faster, suggesting that spatial cueing is affected in two distinct manners. On the one hand, visual attention is directed at the information unit of the diagram, which consists of three nodes connected by two directional line segments. When one of the nodes is cued, the effect automatically spreads to other nodes in the same unit. On the other hand, visual attention is directed at a particular level in the hierarchy; this is guided by geometrical features such as collinearity. However, as shown in the right panel in Fig. 3, when the target appears at the node of a different unit, the node for which the level is higher than the cue is detected faster. These results suggest that the effect of geometrical collinearity is contingent on the information unit to which the node is assigned. If the target node does not belong to the same information unit as the cue node, the higher level is detected faster, indicating that the reading process takes place from a particular direction in the hierarchy



**Fig. 3.** Examples of the stimulus for which the detection of the targets (filled squares) was relatively rapid (the cues are the squares with brighter contours).

(“top bias”, [9]). In order to resolve this confounding, the follow-up experiment was performed by using the same stimuli in the inverted position. According to my unpublished data from this follow-up experiment, the simple effect of level for the interaction was significant only for the same component condition. This result suggested that when the top of hierarchy in the diagram is consistent with that of a visual scene, the viewer receives attentional benefits both from the component of the diagram and the environmental upright. This attentional benefit from the environmental upright eliminates when the hierarchical diagram is presented in the inverted position, but the effect of the component remains.

In sum, the following three different types of spatial features are used in reading a hierarchical diagram to implement an efficient process: belongingness to the component, geometrical collinearity, and top bias. The usefulness of these features is based on the conventional knowledge of hierarchical diagrams that has been acquired throughout encounters with them in daily situations.

### 3 Conclusion

The present study examined how the global structure and local components of a hierarchical diagram influence the cueing effect on a particular node in the diagram. How a particular diagram is constructed and comprehended depends on the conventional knowledge possessed by the observer; the results suggest that different types of perceptual features influence the orientation of visual attention to the encountered diagram. These results might also provide a clue to the appropriate design for diagrams used in a variety of situations in daily life.

The use of conventional knowledge requires a certain amount of processing resources, such as working memory. Hence, individual differences in working memory might be related to performance in a comprehension task. Furthermore,



it has been reported that individuals with autism spectrum disorder require more time for global processing [10]. How such individual differences affect performance in a diagram task might contribute to the concept of universal design, and should therefore be examined in a future study.

**Acknowledgments.** I would like to thank Naoko Takada for conducting the experiment. This work was supported by JSPS KAKENHI Grant Numbers 15K00215 and 15H02716.

## References

1. Novick, L.R., Hurley, S.M.: To matrix, network, or hierarchy: that is the question. *Cogn. Psychol.* **42**(2), 158–216 (2001)
2. Navon, D.: Forest before trees: the precedence of global features in visual perception. *Cogn. Psychol.* **9**(3), 353–383 (1977)
3. Reed, S.K.: *Thinking Visually*. Psychology Press, New York (2010)
4. Shah, P., Freedman, E.G., Vekiri, I.: The comprehension of quantitative information in graphical displays. In: Shah, P., Miyake, A. (eds.) *The Cambridge Handbook of Visuospatial Thinking*, New York, NY, pp. 426–476. Cambridge University Press (2005)
5. Trapp, S., Bar, M.: Prediction, context, and competition in visual recognition. *Ann. N. Y. Acad. Sci.* **1339**(1), 190–198 (2015)
6. Waechter, S., Besner, D., Stolz, J.A.: Basic processes in reading: spatial attention as a necessary preliminary to orthographic and semantic processing. *Vis. Cogn.* **19**(2), 171–202 (2011)
7. Egly, R., Driver, J., Rafal, R.D.: Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol. Gen.* **123**(2), 161–177 (1994)
8. Rosch, E., Mervis, C.B., Gray, W., Johnson, D.M., Boyes-Braem, P.: Basic objects in natural categories. *Cogn. Psychol.* **8**(3), 382–439 (1976)
9. Chambers, K.W., McBeath, M.K., Schiano, D.J., Metz, E.G.: Tops are more salient than bottoms. *Percept. Psychophys.* **61**(4), 625–635 (1999)
10. Van der Hallen, R., Evers, K., Brewaeys, K., Van den Noortgate, W., Wagemans, J.: Global processing takes time: a meta-analysis on local-global visual processing in ASD. *Psychol. Bull.* **141**(3), 549–573 (2015)