

Is size perception based on monocular distance cues computed automatically?

LIAT GOLDFARB and JOSEPH TZELGOV

Ben-Gurion University of the Negev, Beer Sheva, Israel

The study reported here examined whether size perception based on monocular distance cues is computed automatically. Participants were presented with a picture containing distance cues, which was superimposed with a pair of digits differing in numerical value. One digit was presented so as to be perceived as closer than the other. The digits were of similar physical size but differed in their perceptual size. The participants' task was to decide which digit was numerically larger. It was found that the decision took longer and resulted in more errors when the perceptual size of the numerically larger digit was smaller than the perceptual size of the numerically smaller digit. These results show that perceived size affects performance in a task that does not require size or distance computation. Hence, for the first time, there is empirical support for the working assumption of the visual perception approach that size perception based on monocular distance cues is computed automatically.

When operating in an environment, one constantly generates a visual world around oneself. As a part of this, one assesses the distance between objects on the basis of both binocular and monocular distance cues. When looking at the environment or at a picture with one eye closed, one can use only monocular (pictorial) cues to estimate this distance. Distance cues in the vision plane, such as height, perspective, and so on, affect the perceived size of objects that are part of the visual world. Thus, although the retinal size of an object grows smaller as it moves away, one does not perceive the object as actually shrinking when it is distant. This phenomenon, known as *size constancy*, is based on perceiving an object as being more distant (see, e.g., Gibson, 1979; Norman, 1980; Rock, 1983) even when one is unaware of the distance cues (Moore & Egeth, 1997).

There are two main approaches to distance perception. According to the *ecological approach* (Gibson, 1979), generating the size and distance relations among objects results directly from the exploration of visual space. By contrast, the *constructivist approach* argues that the perceived size is computed on the basis of distance cues (see, e.g., Epstein, 1982; Rock, 1983, 1997). Nevertheless, both approaches have some similar features. Gibson claims that

both size and distance are perceived directly. The old theory that the perceiver allows for the distance in perceiving the size of something is unnecessary. . . . The perceptual system must abstract the invariants. The former process seems to be simpler than the latter, more nearly automatic.

Michaels and Carello (1981) interpreted Gibson as indicating that there is no need for active tuning of the perceptual system because it is self-tuning. Epstein, in his constructivist approach, claimed that "the percept in question will have been removed from direct control by information in stimulation," meaning that a stimulus dimension is altered by another stimulus dimension. Similarly, Rock (1997) claimed that "perception is based on prior perception, implying a perception-perception chain of causality." Hence, the working assumption of both approaches to visual perception is that perceiving size on the basis of monocular distance cues is automatic in the sense of not requiring intention and being performed even when not part of the task requirements. However, up until now, there has been no empirical support for that assumption.

Several definitions of automatic processes have been suggested (e.g., Logan, 1992; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). According to Bargh (1989, 1992), the ability of a process to run to completion without conscious monitoring is the single feature common to all automatic processes that characterize human behavior. Accordingly, Tzelgov (1997) proposed to define a process as *automatic* if it could be executed even when it was not part of the task requirement. Hence, he suggested that the best method for diagnosing automaticity in this sense—that is, as processing without monitoring—would be to use Stroop-like phenomena in which the process in question could be shown to affect performance in a task while not being part of the task. Thus the Stroop effect (see, e.g., MacLeod, 1991; Stroop, 1935), in which longer latencies occur for reporting the ink color of a color name in incongruent cases (e.g., BLUE written in red) than in congruent cases (e.g., RED written in red), indicates that the reading of words is automatic in this sense.

Correspondence concerning this article should be sent to L. Goldfarb, Department of Behavioral Sciences, P.O. Box 653, Ben-Gurion University of the Negev, Beer Sheva, Israel 84105 (e-mail: liatgol@bgumail.bgu.ac.il).

It is well documented that under different task sets or other specific conditions, Stroop or Stroop-like effects can be reduced or even eliminated (see, e.g., Bauer & Besner, 1997; Durgin, 2000; Raz, Shapiro, Fan, & Posner, 2002). Yet such effects do diagnose automaticity in the sense of processing without monitoring (e.g., Tzelgov, Yehene, Kotler, & Alon, 2000).

A Stroop-like phenomenon known as the *size congruity effect* is frequently used to study the automaticity of size processing. This effect is obtained when two digits that differ in numerical value and in physical size are presented and participants are asked to decide which digit is numerically larger. It was found that reaction time (RT) needed for processing size for congruent pairs (i.e., the digit that was larger in its numerical value was also larger in its physical size; e.g., 1 9) was faster than for incongruent pairs (i.e., the digit that was larger in its numerical value was smaller in its physical size; e.g., 1 9) (see, e.g., Besner & Coltheart, 1979; Henik & Tzelgov, 1982).

In the present work, we were interested in the automaticity of computing perceived size on the basis of monocular distance cues. We used a technique similar to that used by Robertson and Kim (1999) to investigate the effects of perceived space on attentional cuing. Thus, we superimposed pairs of digits differing in numerical value, but of similar physical size, on a picture containing distance cues. One digit was presented as being closer than the other. Note that, although the physical size of the digits in each pair was similar, they differed in their perceptual size on the basis of distance cues. Thus, in Figure 1, the two digits presented are of similar physical size, yet the digit perceived as farther away is also perceived as larger.

The participants' task was to decide which digit was numerically larger while ignoring the picture and the size of the digit. If, as we hypothesized, the perceived size is automatically computed on the basis of monocular distance cues, the size congruity effect should be obtained even when the digits presented for numerical comparisons are of similar physical size.

METHOD

Participants

Twelve students participated in the experiment in partial fulfillment of course requirements.

Stimuli

At the center of a computer screen, a picture was presented that maximized perceived distance by using a large number of distance cues (see, e.g., Berbaum, Tharp, & Mroczek, 1983; Bruno & Cutting, 1988). As can be seen in Figure 1, the depth cues employed were (1) relative size of familiar objects (the size of the man who is farther away is physically smaller than the physical size of the woman), (2) linear perspective (the "parallel" lines of the carpet start to merge when they appear to be farther in the distance), and (3) height on the plane (the child's base is higher than the man's base, which in turn is higher than the woman's base, hence we should conclude that the child is farther away than the man, who is farther away than the woman). In addition, the picture contains the



Figure 1. Version 1 of the picture that we used, superimposed with the digits 1 and 9. Here we can see an example of the congruent condition, since the digit 9 is larger than 1 in its numerical value and in its perceptual size. The picture is from *Psychology* (5th ed., p. 191), by D. G. Myers, 1998, New York: Worth. Copyright 1998 by Freeman and Company/Worth Publishers. Used with permission.

Müller-Lyer illusion, in which the length of a line that is part of the arrow seems longer when its upper and lower arrow heads are directed outside, rather than inside. The parts of the picture that generated the Müller-Lyer illusion are the arrow with its heads pointing inside part of the ticket window and the arrow with its heads pointing outside part of the inside room corner. Although in both cases the lines of the arrows are similar in physical size, the perceived size of the line of the ticket window is smaller. The pairs of digits appear in these two places. The picture was presented in two versions: In the first version, the far corner was on the left side of the screen (as in Figure 1). The second version was the mirror image of the first version so that the far corner was on the right side of the screen.

In every block, the digits 1–9, excluding 5, were coupled to create the 28 possible pair combinations. Each combination was presented six times per side in each of the two versions of the picture. Because there were two blocks in the experiment, the total number of experimental trials was 1,344 (28 digit pairs \times 2 sides \times 6 presentations \times 2 picture versions \times 2 blocks).

Procedure

The experiment was run on a Compaq computer with an Intel Pentium III central processor. The stimuli were presented on a Compaq S510 monitor. The participants were run individually and sat approximately 80 cm from the monitor. At the beginning of the experiment, the participants were instructed to select the numerically larger digit while ignoring the picture and the size of the digits. They were instructed to press the "p" button on the keyboard if the numerically larger digit appeared on the right side of the screen and to press the "q" button if it appeared on the left. They were asked to respond as quickly as possible without making mistakes. Eight practice trials preceded the experimental trials.

Each trial started with a fixation point—that is, a white "+" sign that appeared for 300 msec at the center of a black screen. Five hundred msec after its disappearance, a picture with a pair of superim-

posed digits similar in physical size¹ appeared at the center of the screen for 3,500 msec or until the participant responded. The computer measured RTs in milliseconds from the onset of the stimulus to the onset of the participant's manual response. After the participant's response, the stimulus disappeared and the next trial began. If the participant's response was incorrect, an audio feedback was sounded for 500 msec, and only then did the next trial begin.

RESULTS AND DISCUSSION

Mean RTs of correct responses to the target and mean error percentages were calculated for each participant for congruent pairs (i.e., the numerically larger digit was also perceptually larger) and for incongruent pairs (i.e., the numerically larger digit was perceptually smaller). RTs faster than 250 msec were not analyzed (only one trial was faster than 250 msec). A paired-samples *t* test was applied on the RT data and on the error percentage data.

The means of the RTs and the error rates in the various conditions are presented in Table 1. As can be seen, the responses in the congruent condition were faster than those in the incongruent condition [$t(11) = 4.61, p < .001$], and errors in the congruent condition were fewer than those in the incongruent condition [$t(11) = 3.35, p < .01$].²

Our results show a size congruity effect, indicating computations of perceived size based on monocular distance cues, even when such computations were not part of the task requirements, and thereby verifying the automaticity of these computations. Thus, whether the perception of size based on monocular distance cues is a direct result of an exploration of visual space as stated by the ecological view (see, e.g., Gibson, 1979), or whether it results from a deductive inference as claimed by the constructivist approach (Rock, 1983, 1997), our results show that the process of utilizing monocular cues and generating size perception is automatic and happens even when it does not contribute to our performance in the world.

Using similar logic, Moore and Egeth (1997) have shown that perceptual grouping does not require attention. They have shown that when pairs of lines of equal length are tachistoscopically presented for length comparison, the appearance of irrelevant patterns of dots resulted in the Ponzo (Experiments 1 and 2) or the Müller-Lyer illusion. From the viewpoint of the present analysis, Moore and Egeth's result indicates that perceptual grouping is automatic in the sense that it does not require attention. We are not arguing that the perception of distance on the basis of monocular cues is preattentive. We

do believe, however, that both perceptual grouping and the computation of distance on the basis of pictorial cues reflect the automaticity of visual perception.

REFERENCES

- BARGH, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3-51). New York: Guilford.
- BARGH, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology*, **105**, 181-199.
- BAUER, B., & BESNER, D. (1997). Mental set as a determinant of processing in the Stroop task. *Canadian Journal of Experimental Psychology*, **51**, 61-68.
- BERBAUM, K., THARP, D., & MROCZEK, K. (1983). Depth perception of surfaces in pictures: Looking for conventions of depiction in Pandora's box. *Perception*, **12**, 5-20.
- BESNER, D., & COLTHEART, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia*, **17**, 467-472.
- BRUNO, N., & CUTTING, J. E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology: General*, **117**, 161-170.
- DURGIN, F. H. (2000). The reverse Stroop effect. *Psychonomic Bulletin & Review*, **7**, 121-125.
- EPSTEIN, W. (1982). Percept-percept coupling. *Perception*, **11**, 75-83.
- GIBSON, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- HENIK, A., & TZELGOV, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, **10**, 389-395.
- LOGAN, G. D. (1992). Attention and preattention in theories of automaticity. *American Journal of Psychology*, **105**, 317-339.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.
- MICHAELS, C. F., & CARELLO, C. (1981). *Direct perception*. Englewood Cliffs, NJ: Prentice Hall.
- MOORE, C., & EGETH, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 339-352.
- NORMAN, J. (1980). Direct and indirect perception of size. *Perception & Psychophysics*, **28**, 306-314.
- POSNER, M. I., & SNYDER, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.
- RAZ, A., SHAPIRO, T., FAN, J., & POSNER, M. I. (2002). Hypnotic suggestion and the modulation of Stroop interference. *Archives of General Psychiatry*, **59**, 1155-1161.
- ROBERTSON, L. C., & KIM, M. S. (1999). Effects of perceived space on spatial attention. *Psychological Science*, **10**, 76-79.
- ROCK, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
- ROCK, I. (1997). *Indirect perception*. Cambridge, MA: MIT Press.
- SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, **84**, 127-190.
- STROOP, J. R. (1935). Studies in interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-661.
- TZELGOV, J. (1997). Specifying the relations between automaticity and consciousness: A theoretical note. *Consciousness & Cognition*, **6**, 441-451.
- TZELGOV, J., YEHENE, V., KOTLER, L., & ALON, A. (2000). Automatizing algorithmic processing: The learning of linear ordering relations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 103-120.

NOTES

1. For all practical purposes, the physical sizes of the presented digits were identical. The physical size of the digits was 4.64° with a max-

Table 1

Mean Reaction Times (RT, in Milliseconds) and Percent Error Rates in the Congruent and Incongruent Conditions

Condition	RT	%ER
Congruent	678	1.45
Incongruent	714	3.46

imum deviation of 0.03° between the different digits. This deviation emerged because the different digits varied in shape. Given that the exact same pair of numbers appeared for the congruent and the incongruent conditions, this deviation was irrelevant to the effect. Moreover, the digits 2, 3, 6, 8, and 9 were exactly equal in physical size. Hence, we performed another analysis that included only these digits. The congruency effect also emerged in this case [$t(11) = 4.58, p < .001$].

2. The pairs of digits were superimposed on a single picture. Hence, we examined whether practice influenced the congruity effect reported.

A two-way analysis of variance with congruency and block as within-participants factors was applied on the RT data. The interaction between congruency and block was not significant ($F < 1$). Consequently, it seems that practice on only one picture did not affect the magnitude of the congruency effect.

(Manuscript received June 7, 2004;
revision accepted for publication November 29, 2004.)