Title
Two Views are Better than One: Epistemic Actions May Prime

Permalink
https://escholarship.org/uc/item/70c0d6x2

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

Authors
Maglio, Paul P.
Wenger, Michael J.

Publication Date
2000

Peer reviewed

Two Views are Better than One:
Epistemic Actions May Prime

Paul P. Maglio
IBM Almaden Research Center
San Jose, California
pmaglio@almaden.ibm.com

Michael J. Wenger
Department of Psychology
University of Notre Dame
mwenger1@nd.edu

Abstract

Epistemic actions are physical actions people take more to simplify their internal problem-solving processes than to bring themselves closer to an external goal state. In the video game Tetris, for instance, players routinely over-rotate falling shapes, presumably to make recognition or placement decisions faster or less error-prone. Along these lines, an experimental study was performed to test the hypothesis that it is easier to recognize a two-dimensional shape if it is presented in two different orientations than if it is presented in only one. In particular, we tested whether performance on a shape-based video game task was facilitated by multiple views of a shape, and whether game performance (an indirect test of memory) differed from a direct test of memory for previously presented shapes. Results show that indeed task performance is both faster and more accurate when participants see two views of a shape than when they see one, but that more than two views do not improve performance further. In addition, multiple views lead to faster performance on the video game than on the memory test, but only in the earliest stages of training. We conclude that Tetris players may rotate falling shapes manually to see the shapes in more than one orientation, which leads to faster and more accurate placement decisions.

Introduction

Studies of people playing the video game Tetris have shown players often take actions in the external environment that are not strictly necessary but that serve to simplify or speed up internal cognitive or perceptual operations (Kirsh & Maglio, 1994; Maglio, 1995; Maglio & Kirsh, 1996). Playing Tetris involves maneuvering falling two-dimensional shapes into specific arrangements on the computer screen (see Figure 1). It was found that even as players become faster with practice, they also tend to over-rotate falling shapes, leading to backtracking in the task environment as these over-rotations are corrected. To make sense of this backtracking, Kirsh and Maglio (1994) argued that sometimes physical rotation can serve the same purpose as mental rotation, effectively offloading mental computation onto the physical world (for other examples, see Clark, 1997; Kirsh, 1995; Maglio, Matlock, Raphaely, Chernicky & Kirsh, 1999). Such physical actions—taken to simplify internal cognitive computation rather than to move closer to the external goal state—are called *epistemic actions*.

Recent work suggests that mental rotation and physical rotation share at least some internal processes (e.g.,

Wexler, Kosslyn & Berthoz, 1998; Wexler & McIntyre, 1997; Wohlschlager & Wohlschlager, 1998). Specifically, physically rotating objects can be shown to facilitate or to inhibit mental rotation under certain conditions. The epistemic function of physical rotation in Tetris, therefore, might be far more complex than is suggested by the simple idea that physical rotation can substitute for mental rotation. In fact, Kirsh and Maglio (1994) speculated that physical rotation might serve the epistemic function of cueing retrieval. Because physically rotating a game piece (which we call a *zoid*) in Tetris provides the player two views of it (i.e., in each of two orthogonal orientations), it is possible that seeing two views makes retrieval of relevant information easier than does seeing just one. This idea makes computational sense; for example, if one conceives of memory in terms of an attractor space, such as a Boltzman machine, the first presentation of the shape is like placing the system near the top of the energy sink that represents the target shape in memory, and the second pushes the system closer to this attractor.

Of course, if shape recognition is orientation-dependent (Tarr & Pinker, 1989; Tarr, 1995; Ullman, 1989), we would not expect multiple views of a single shape to speed up recognition. However, it has been shown that shape identification can be facilitated when primed with orientations different from the target orientation (Cooper, Schacter, Ballesteros & Moore, 1992; Srinivas, 1995). Moreover, numerosity judgments can be facilitated even when test stimuli are not presented at the same orientation as the originally learned patterns (Lassaline & Logan, 1993), suggesting memory for the pattern may not require that the retrieval cue be specifically oriented.

That an epistemic action might cue retrieval raises the possibility that such cueing might be limited to specific types of retrieval demands. In particular, the effects of cueing might depend on whether the task requires *direct* or *indirect* access to memory information. Demands for retrieval while playing Tetris can be thought of as *indirect* tests of memory in that they allow for effects of prior experience to be expressed without requiring explicit memory for the original experience (e.g., Richardson-Klavehn & Bjork, 1988). Tasks requiring explicit memory for the original event—such as old/new recognition or recall—are referred to as *direct* tests of memory. Previous work has shown that direct and indi-
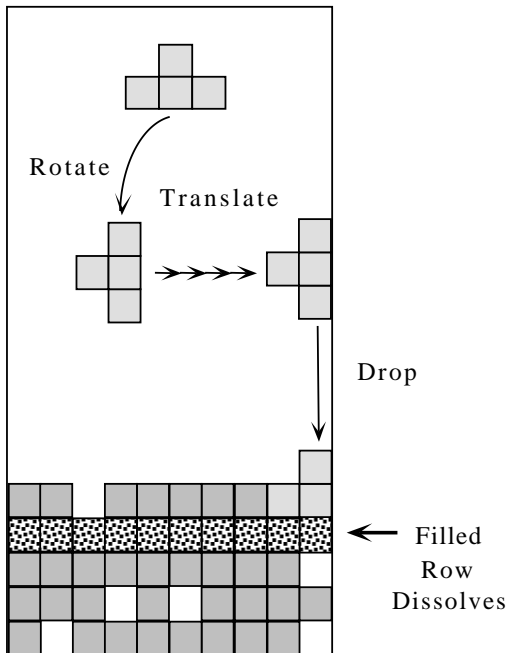
Figure 1: In Tetris, two-dimensional shapes fall one a time from the top of the screen, eventually landing on the bottom or on top of shapes that have already landed. There are seven shapes, which we call *zoids*— ⊟ , ⊞ , ⊞ , ⊞ , ⊞ , ⊟ , ⊟ . As a zoid falls, it can be rotated, and moved to the right or left. The object of the game is to fill rows of squares all the way across the screen. When a row is completely filled, it dissolves and all partially filled rows above it move down. The game ends when unfilled rows pile up to the top, blocking new zoids from falling.

rect tests of memory are differentially sensitive to characteristics such as orientation, object symmetry, and other physical aspects of visual objects (Srinivas, 1995, 1996; Srinivas & Schwoebel, 1998). Thus, in the experiment presented here, we used both direct and indirect assessments of memory to determine how effective previews are under different retrieval demands. In addition, because the effectiveness of memory cues generally depends on the time that elapses between presentation of cue and presentation of the item to be retrieved, we investigated the effect of various delays between onset of the first preview and onset of the test zoid by embedding the previews in a sequence of zoids presented prior to test.

In this paper, we empirically test the hypothesis that two different views of a falling zoid are better than one. In addition, we examine whether such a potential benefit might depend on the orientation of the preview relative to the zoid that must be placed, and whether these previews facilitate zoid recognition and Tetris performance.

Method

To test whether two views of a falling zoid leads to faster or more accurate performance in Tetris than does one, we created a controlled experimental situation that shared many attributes with the game of Tetris but that allowed fine-grained control over the parameters of interest. In our experimental set up, a Tetris configuration (i.e., a Tetris board and zoid floating above it) is preceded either by none, one, or two previews of the zoid in either the same or different orientations (see Figure 2). The participant's job is to quickly and accurately determine whether the zoid fits snugly on the board. Thus, the task creates situations similar to those faced by Tetris players during an actual game, and also requires responses similar to those required of players during an actual game.

Participants spent three days (one hour each day) playing this experimental version of Tetris. Separate groups of participants were required either (a) to make judgments about whether a target zoid fit in an accompanying board (the indirect test), or (b) to make this judgment *and* indicate whether they remembered seeing the test zoid in the set of zoids that were presented prior to the target (the direct test). Between 0 and 2 previews of the target zoid were presented in a sequence of zoids prior to the target, and the orientation of these previews (when present) varied relative to the target. As noted, by placing the previews in a sequence of events prior to the test, we were able to manipulate the interval over which the preview would have to be retained in memory.

Participants

A total of 30 participants were recruited from psychology courses and participated voluntarily in exchange for course credit. All participants reported normal or corrected-to-normal vision and unencumbered use of both hands.

Design

The experiment was conducted as a 3 (number of previews: 0, 1, 2) × 3 (orientation of the first preview relative to the target zoid: same, clockwise rotation of 90°, counter-clockwise rotation of 90°) × 3 (retention interval between first preview and target zoid, in frames: 0, 1, 2) × 3 (zoid type: ⊞ , ⊟ , ⊞) × 2 (status of target zoid relative to the board: fit, not fit) × 3 (day of testing: 1, 2, 3) × 2 (type of memory judgment at test: direct, indirect) mixed factorial design. All factors except type of zoid and type of memory judgment were manipulated within participants.

Materials

All zoids and boards were constructed from 20×20 pixel squares. Squares were outlined by light gray lines, 1 pixel in width, and were filled in solid black. The background for all displays was solid black as well. All zoid types were composed of four blocks. All receptor boards were six blocks in height and width. Four receptor types were defined for each zoid type, corresponding to four ways in which the zoid could be snugly placed. Each receptor type was used with equal frequency. Materials were displayed on a 33 cm VGA monitor controlled by a PC-compatible microcomputer. Onset and offset of each display was synchronized to the vertical scan of the
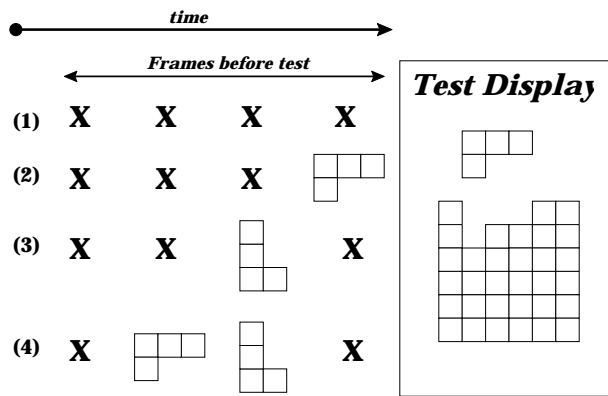
Figure 2: A schematic representation of some of the events in four frames prior to a test display in a single trial. The Xs indicate non-target zoids. (1) The events in a 0-preview trial. (2) The events in a 1-preview trial, with no retention interval (0 frames) between the preview and the test display. (3) The events in a 1-preview trial with a 1-frame retention interval between the preview and the test display. Here the preview is rotated 90° counter-clockwise relative to the test display. (4) The events in a 2-preview trial with a 2-frame retention interval between the *first* preview and the test display. Here the first preview has the same orientation as the test display, while the second preview is rotated 90° counter-clockwise relative to the test display.

monitor. The standard PC keyboard was used to collect and time (to $\pm 1ms$) participant responses.

Procedure

Participants were tested on three consecutive days, at approximately the same time each day, with each session lasting approximately 1 hour. All sessions were conducted in a darkened room, with participants seated at an unconstrained distance from the monitor, and began with a five min period for dark adaptation. Participants were told that, on each trial, they would see a sequence of zoids, presented very rapidly. At some random point in this sequence, they would see a combination of a zoid and a receptor board, and would need to make one of two types of responses, depending on whether they were in the indirect or direct memory test condition.

In the indirect condition, participants simply had to decide whether the presented piece would fit snugly into the board. Participants responded in the affirmative using the index finger of their dominant hand, and in the negative using the index finger of their non-dominant hand, pressing either the "z" or "/" keys on the lower row of the PC keyboard. In the direct condition, participants had to indicate with a single key-press both their judgment about whether the presented piece fit snugly in the board and their memory for any occurrence of the

test piece (in any orientation) in the sequence of pieces that preceded the target piece. Participants responded with the index finger of their dominant hand if the target piece fit and they remembered seeing this piece in the preceding sequence, with the middle finger of their dominant hand if the target piece fit and they did not remember seeing this piece in the preceding sequence, and with the index finger of their non-dominant hand if the piece did not fit.[1] Speed and accuracy were equally emphasized.

Each trial began with the presentation of between one and eight zoids ("non-target zoids") designed to be distinct from the target zoid assigned to the participant. The actual number of these non-target zoids shown was randomly determined for each trial. Each non-target zoid was presented for 250 ms and then replaced by the next non-target zoid; the non-target zoids in this sequence did not repeat (i.e., all were unique). Following this, four zoids (between 0 and 2 target zoids, and between 2 and 4 non-target zoids) were presented for 250 ms each. After the last of these were presented, a target zoid and a receptor board were presented for 250 ms. Following the participant's response, a tone was briefly sounded (100 ms) indicating a correct (880 Hz) or incorrect (440 Hz) response.

A total of 480 trials were presented in each session. Participants were allowed short breaks after every 80 trials. Feedback on overall accuracy and mean response time was provided at the end of each session.

Results

First, we asked whether having one preview improved performance over having no previews, and found a pronounced effect in both accuracy and response time (RT). When participants were presented with a single preview, the resulting level of accuracy was significantly higher (0.86) than when they were not presented with a preview (0.53), $t_{(1,59)} = 33.85, p < 0.001$. Similarly, when participants were presented with a single preview, the resulting RTs were significantly shorter (869 ms) than when they did not see a preview (1791 ms), $t_{(1,59)} = 2.01, p < 0.05$.

Given that providing a preview had an effect on performance, we moved on to determining whether having *more than one* preview had an additional effect, and whether the provision of previews interacted with our other experimental factors. Our analysis of the accuracy data indicated that zoid, number of previews (1 vs. 2), and retention interval all failed to have an effect on accuracy (all $Fs < 1.00$). However, test type did have a significant impact on performance, with participants in the direct test condition performing at a higher level of accuracy (0.95) than participants in the indirect condition (0.88), $F_{(1,25)} = 4.59$, MSE = 0.05. Orientation of the prime exerted a statistically significant effect on accuracy, $F_{(1,25)} = 4.01$, MSE = 0.01, but the magnitude of the difference between the previews presented in the

---

[1] We did not ask for a memory judgment on trials in which the piece was judged not to fit, as our primary concern was with the effects of previews on accurate placement of pieces in the board.
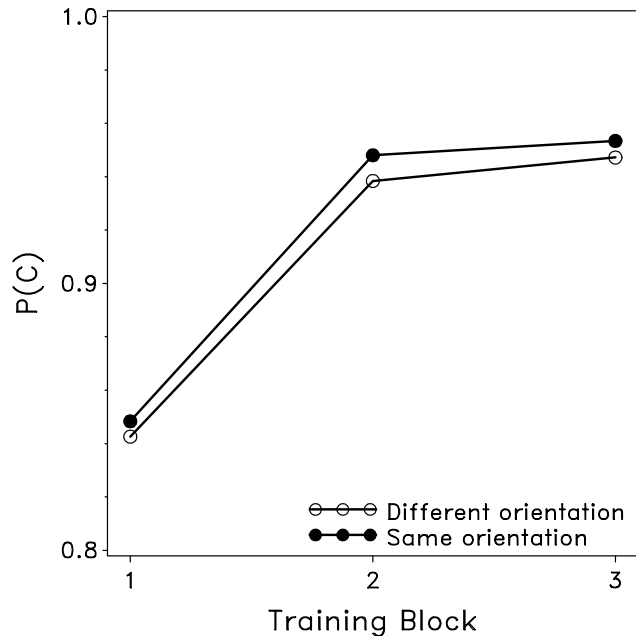
Figure 3: Effects of orientation of preview and block on accuracy. Practice affects the probability of making a correct response. However, whether the zoid was previewed in the same orientation or in a different orientation (as the test zoid) does not affect the probability of making a correct response.
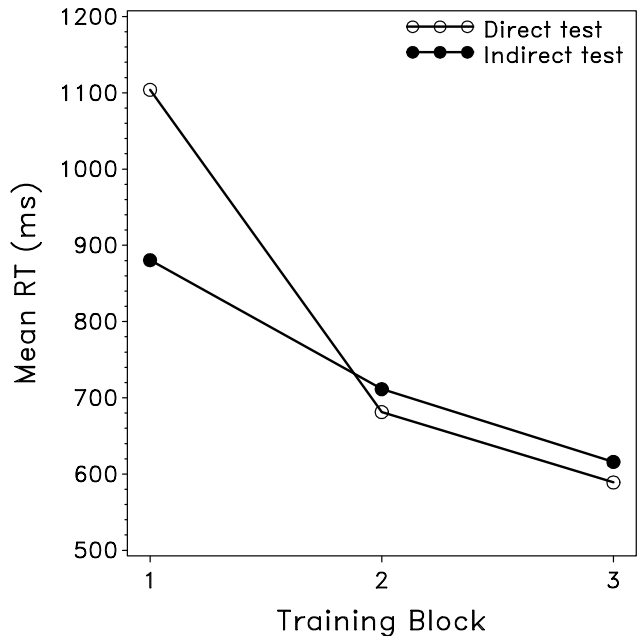

Figure 4: Effects of test type and block on mean RT. Participants in the indirect test condition (i.e., deciding whether the zoid fits snugly) respond faster than participants in the direct condition (i.e., deciding whether the zoid fits *and* whether the zoid had been previewed) only on the first day of practice.

same orientation (0.92) and those presented in a different orientation (0.91) suggests that the difference may not be meaningful. Exploration of these data across blocks of experience (see Figure 3) suggests that the difference between the two forms of preview was induced by the fact that performance with previews in a different orientation did not improve quite as quickly from the first to the second training block as did performance with previews in the same orientation, though this interaction was *not* significant. Finally, as expected, performance improved consistently across blocks, $F_{(2,50)} = 6.67$, MSE = 0.03, as can be seen in Figure 3.

Analysis of the RT data indicated that test type, zoid, number of previews, orientation of the preview, and retention interval all failed to affect the speed of responding (all $Fs < 1.00$). Although RTs consistently improved across the experiment, $F_{(2,50)} = 57.56$, MSE = 44847.84, the form of improvement was dependent on test type (direct vs. indirect), $F_{(2,50)} = 7.03$, MSE = 44847.84. As shown in Figure 4, the direct test condition (which required *two* response judgments) was slower than the indirect test condition (which required one response judgment), but only in the first block of trials.

Discussion

Our results show that if participants are presented with two views (i.e., one preview) of the falling zoid (a two-dimensional shape), response times are faster than if only a single view (i.e., no previews) is presented. This sup-

ports our hypothesis that *two views are better than one*. Nevertheless, it was a bit surprising to find that three views provide no advantage over two views. In terms of the simple Boltzman machine model mentioned previously, this would mean that the second view of the zoid pushes the system so close to the attractor that it is trapped, and so the third view is rendered irrelevant. Alternatively, the effect of the first preview might be to accelerate the system toward the attractor state to such an extent that a second preview provides no appreciable additional acceleration.

Note that response time was speeded up by a preview in any of the three orientations relative to the test zoid. The benefit was not restricted to previews that shared orientation with the test display. This finding is consistent with priming studies in which it was found that a prime need not be presented in the same orientation as the target to facilitate recognition or identification (e.g., Cooper, Schacter, Ballesteros & Moore, 1992; Srinivas, 1995). It is surprising, however, to find that different orientations prime just as strongly as the test orientation does. One possible explanation is that participants have stored multiple views of the zoids and so seeing one view is just as good as seeing another (Tarr & Pinker, 1989).

The only difference between the direct and indirect tests of memory was observed on the first day of training, and restricted to the latency data. On the first day, participants in the direct test condition required more time

to respond than did the participants in the indirect test condition. This difference may be easily accounted for by the fact that participants in the direct test condition had to make two response decisions and choose among three response alternatives. The *lack* of a difference in either accuracy or latency as a function of memory test suggests that the benefits obtained by having a preview do not depend on the manner in which memory for that preview is assessed.

Returning to the idea of epistemic action in Tetris, these data suggest that by rotating the falling zoids, players may be able to effectively cue themselves, enabling quicker responses in a Tetris situation. Previous research has established various ways in which Tetris players take actions for their epistemic effects (Kirsh & Maglio, 1994; Maglio, 1995; Maglio & Kirsh, 1996). The data reported here show that a preview of the falling zoid at least speeds up performance on a Tetris-like task, but the hypothesis that Tetris players over-rotate zoids in order to speed up performance is not directly tested. It remains to be seen whether actually taking the action of orienting the preview (i.e., physically rotating the falling shape) is a critical component of performance, independent of the presentation of the preview itself.

In the end, we can conclude that two sequentially presented views of the falling zoid lead to faster and more accurate performance than a single view of the falling zoid. In addition, it appears that having this single preview is sufficient to boost performance to something of a limit, as more than one preview adds little if any additional help. It also appears that the benefit of the preview is robust across the retention intervals considered here. Thus, if players are able to use rotations to self-cue, they may be able to get all they need from a single rotation, even one that is somewhat separated in time from the eventual judgment. The payoff associated with a small number of additional steps more than compensates for the temporal and physical costs of executing additional steps. The epistemic functions of physical rotations in Tetris, then, might not be merely to substitute for mental rotation or to provide a visual means for matching the contour of the board with contour of the falling shape, but also to cue or prime retrieval from memory of information associated with the falling shape, enabling faster recognition and faster placement decisions.

Acknowledgments

References

Clark, A. (1997). *Being there: Putting body, brain, and world together again*. Cambridge, MA: MIT Press.

Cooper, L. A., Schacter, D. L., Ballesteros, S., & Moore, C. (1992). Priming and recognition of transformed three-dimensonal objects: Effects of size and reflection. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*, 43–57.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence, 73*, 31–68.

Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18*, 513–549.

Lassaline, M. E. & Logan, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning Memory and Cognition, 19*.

Maglio, P. P. (1995). *The computational basis of interactive skill*. Doctoral dissertation, University of California, San Diego.

Maglio, P. P. & Kirsh, D. (1996). Epistemic action increases with skill. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, pages 391–396, Mahwah, NJ. Lawrence Erlbaum.

Maglio, P. P., Matlock, T., Raphaely, D., Chernicky, B., & Kirsh, D. (1999). Interactive skill in Scrabble. In *Proceedings of the Twenty-first Annual Conference of the Cognitive Science Society*, pages 326–330, Mahwah, NJ. Lawrence Erlbaum.

Richardson-Klavehn, A. & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39*, 475–543.

Srinivas, K. (1995). Representation of rotated objects in explicit and implicit memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*, 1019–1036.

Srinivas, K. (1996). Contrast and illumination effects on explicit and implicit measures of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1123–1135.

Srinivas, K. & Schwoebel, J. (1998). Generalization to novel views from view combination. *Memory & Cognition, 26*, 768–779.

Tarr, M. & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*, 233–282.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin and Review, 2*, 55–82.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32*, 193–254.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition, 68*, 77–94.

Wexler, M. & McIntyre, J. A. (1997). Is mental rotation a motor act. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, pages 808–813, Mahwah, NJ. Lawrence Erlbaum.

Wohlschlager, A. & Wohlschlager, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 397–412.