UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

On the Potential of Epistemic Actions for Self-Cueing: Multiple Orientations Can Prime 2D Shape Recognition and Use

Permalink

Journal

ISSN

Authors

Maglio, Paul P
Wenger, Michael J

Publication Date

2002

Peer reviewed

On the Potential of Epistemic Actions for Self-Cueing: Multiple Orientations Can Prime 2D Shape Recognition and Use

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

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

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. Consider how when playing the video game Tetris, experts routinely rotate falling two-dimensional shapes more than is necessary to place the shapes. One reason for such apparently unnecessary actions is that they actually help the player make placement decisions. Such actions might facilitate placement decisions if additional previews of the shape afforded by rotating it provide information about the board, particularly when there is no direct perceptual match between the shape and the board at the time of decision. The study presented here tests the hypothesis that several distinct previews of a two-dimensional shape can improve a person's ability to recognize and use that shape when it is not correctly oriented at the time of decision. Results show that indeed task performance and recognition are faster with *two different* orientations than with only one. Thus, it is possible that Tetris players rotate two-dimensional Tetris shapes manually to see them in more than one orientation, as this can lead to faster decisions.

Introduction

People playing the video game Tetris often take actions that are not strictly necessary but that serve to simplify or speed up internal cognitive or perceptual operations (Kirsh & Maglio, 1994; Maglio & Kirsh, 1996). Playing Tetris involves maneuvering falling two-dimensional shapes into specific arrangements on the computer screen (see Figure 1). Even as players become faster with practice, they tend to over-rotate falling shapes, leading to backtracking as these over-rotations are corrected. To make sense of such 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 (see also 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*.

Because shape identification can be facilitated when primed with orientations different from the target orientation (Cooper, Schacter, Ballesteros & Moore, 1992; Srinivas, 1995), and because numerosity judgments can be facilitated even when test stimuli are not presented at the same orientation as the originally learned patterns (Lassaline & Logan, 1993), memory for a target pattern might not require the retrieval cue be specifically oriented. Thus, the epistemic function of physical rotation in Tetris might be far more complex than is suggested by the simple idea that physical rotation can substitute for mental rotation, for instance, serving the function of cueing retrieval (Kirsh & Maglio, 1994). Because physically rotating a Tetris shape (which we call a *zoid*) provides the player two views of it (i.e., in each of two orientations), it is possible that seeing two different views makes retrieval of relevant information easier than does seeing just one. In fact, we found previously that when participants in a Tetris-like task are presented with two views of a zoid, the time taken to decide whether it fits a particular board is faster than when participants are presented with only a single view, but this does not depend on the orientation of the previews relative to one another (Maglio & Wenger, 2000).

There are at least three potential functions of rotation for self-cueing in Tetris. First, seeing the falling zoid in several different orientations may provide helpful information about the board, particularly when there is no direct perceptual match between zoid and board at decision time. That is, if the orientation of the zoid floating above the board does not match the orientation in which the zoid actually fits the board when the player must decide to place it, then having previously seen the zoid in the matching orientation might help the process of mentally matching zoid and board. Such an effect might be the result of having recently seen the zoid in its fitting orientation, basing the decision as to whether the zoid matches the board on memory rather than on mental rotation. Let us call this potential epis-
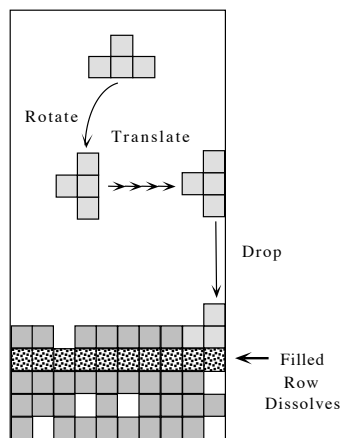
Figure 1: In Tetris, two-dimensional shapes fall one a time from the top of the screen, landing on the bottom or on top of shapes that have already landed. There are seven shapes, or *zoids*—▯, ⊞, ⊞, ⊞, ⊞, ⊞, ⊞. As a zoid falls, it can be rotated, and moved right or left. The object of the game is to fill rows of squares all the way across. Filled rows dissolve and all unfilled rows above move down.

temic function of rotation, the *board-match* function.

A second potential function of rotation for self-cueing might be to provide advance information about the zoid itself, particularly when the several previews coincide with the orientation of the zoid at the time of decision. That is, if the orientation of the zoid floating above the board matches the orientation in which the zoid fits the board when the player must decide to place it, then having seen it previously in that orientation might make recognition easier. In this case, such an effect might be the result of a complex memory retrieval process in which multiple views of a shape lead to faster or more reliable recognition of it (see Maglio & Wenger, 2000). Let us call this potential epistemic function of rotation, the *zoid-retrieval* function.

A third potential function of rotation in Tetris might relate to motor processes rather than to memory or perceptual processes. Because physically rotating objects can facilitate or inhibit mental rotation under certain conditions, it is possible that mental rotation and physical rotation share at least some internal processes (e.g., Wexler, Kosslyn & Berthoz, 1998). Thus, the specific motor act Tetris players take in rotating the falling zoid might serve the purpose of coordinating motor processes with other internal processes to facilitate zoid placement decisions. Let us call this potential epistemic function, the *motor-process* function.

These three epistemic functions of action—the

board-match function, the zoid-retrieval function, and the motor-process function—are not mutually exclusive. All are possible reasons for the over-rotations observed in normal Tetris play. In this paper, we explore only the board-match function. Specifically, we test the hypothesis that seeing *several different* orientations of a falling zoid is better than seeing just one when the final orientation of the zoid does not match the region the zoid fits on the contour of the board. As noted, any such facilitation might result from matching the board to the memory of the previewed zoid rather than mentally rotating the zoid seen at test. Thus, our board-match hypothesis is a kind of memory-retrieval hypothesis.

Retrieval demands 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. Because direct and indirect tests are differentially sensitive to orientation, object symmetry, and other physical aspects of visual objects (Srinivas, 1995; Srinivas & Schwoebel, 1998), the experiment presented here used both direct and indirect assessments of memory to determine how effective previews are under different retrieval demands. 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 also investigated the effect of various delays between final preview and onset of test.

Method

To test whether two orientations of a falling zoid leads to faster performance in Tetris than one orientation does, 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 (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. The participant's job is to quickly and accurately determine at the time of test whether the zoid floating above the board 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. In all cases, in the final Tetris configuration, the zoid and the region it fits on the board contour (if it fits) are oriented
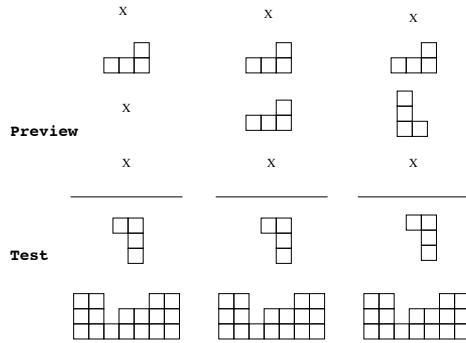
Figure 2: Three trial types used, from left to right: one preview, two previews in the same orientation, and two previews in different orientations. An "X" indicates display of an irrelevant zoid for the trial.
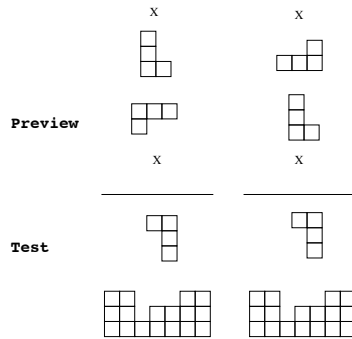

Figure 3: The second preview zoid is oriented properly relative to the board in the trial on the left but not in the one on the right.

differently, meaning there was no perceptual match between zoid and board at test (see Figure 2). In some cases, the last preview was oriented so as to fit snugly on the board contour without rotation, in which case memory for the previewed zoid might facilitate the fit/no-fit decision (see Figure 3).

Participants spent about three hours playing our 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 (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 (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. 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

Twenty-nine participants were recruited from psychology courses and participated voluntarily in exchange for course credit: 15 in the indirect condition, and 14 in the direct condition. All participants reported normal or corrected-to-normal vision.

Design

The experimental design was fairly complicated so as to control as many factors as possible. As described, our main interest was in whether multiple previews of the zoid primed recognition and use better than a single preview when there was no perceptual match between zoid and board at test. In addition, we controlled whether the test zoid fit the test board, whether the preview zoids were in the same or different orientations, the time between preview and test, and whether memory was tested directly (asking whether the test zoid had been previewed) or indirectly (asking only for a fit/no-fit judgment).

More precisely, the experiment was conducted as a 4 (preview type: no previews, one preview, 2 previews same orientation, 2 previews different orientation) × 2 (orientation of the last preview relative to the board: same, different) × 3 (retention interval between last preview and target zoid, in frames: 0, 1, 2) × 2 (zoid type: $\boxplus$, $\boxplus$) × 2 (status of target zoid relative to the board: fit, not fit) × 2 (type of memory judgment at test: direct, indirect) mixed factorial design. All factors except 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 also solid black. All zoid types were composed of four blocks. All boards were six blocks in height and width. Four "fit" boards were defined for each zoid type, corresponding to four ways in which the zoid could be snugly placed. Each such board was used with equal frequency. Materials were displayed on a 33 cm VGA monitor controlled by a PC-compatible computer. Onset and offset of each display was synchronized to the monitor's vertical scan. A standard keyboard was used to collect and time (to $\pm 1 ms$) responses.

Procedure

Participants were tested on two consecutive days, at approximately the same time each day, with each session lasting approximately 90 min. All sessions were conducted in a darkened room, with participants seated 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. The zoids in the sequence would begin falling from a location near the top of the screen: each successive zoid would appear below the one before to create a sequence of falling zoids much as in the Tetris game. Each zoid was present for 250 ms, and each sequence consisted of between five and seven zoids, with the actual number determined randomly (and with equal likelihood) on each trial. At some random point in this sequence, participants would be presented with a combination of a test zoid and board, and would need to make one of two types of responses, depending on whether they were in the indirect or direct memory condition.

In the indirect condition, participants had to decide whether the zoid presented at test would fit snugly into the board. Participants responded in the affirmative using the index finger of the dominant hand, and in the negative using the index finger of the non-dominant hand, pressing either the "z" or "/" keys on the lower row of the keyboard. In the direct condition, participants had to indicate with a single key-press both judgment about whether the presented zoid fit snugly in the board and memory for any occurrence of the test piece (in any orientation) in the sequence that preceded the target. Participants responded with the index finger of the dominant hand if the target piece fit and they remembered seeing this piece in the preceding sequence, with the middle finger of the dominant hand if the target piece fit and they did not remember seeing it in the preceding sequence, and with the index finger of the non-dominant hand if the piece did not fit. Speed and accuracy were emphasized equally.

Results

Note that participants quickly became very good at this task; by the end of the first day, overall error rate was below 3%, indicating a high level of skill. Now, to determine whether primes had an effect, correct reaction times (RT) were analyzed using two 2 (zoid type: ⊞ , ⊟) × 2 (preview: present, absent) × 2 (status of the test zoid: does fit, does not fit) repeated measures ANOVAs, one for each test condition (direct, indirect). In both test conditions, presence of a preview speeded responses (indirect, 556 ms vs. 691 ms, $F_{(1,14)} = 11995.0$, MSE = 23.04; direct, 867 ms vs. 1008 ms, $F_{(1,13)} = 1095.75$, MSE = 254.10). Both conditions also showed faster responses when the test zoid fit relative to when it did not (indirect, 594 ms vs. 653 ms, $F_{(1,14)} = 43.67$, MSE = 1214.50; direct, 890 ms vs. 985 ms, $F_{(1,13)} =$
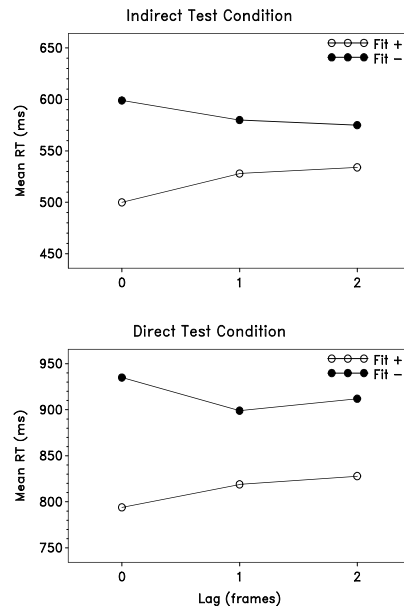

Figure 4: Preview-present trials: interaction of lag and fit status. Fit + indicates trials where the test zoid fit, and Fit -, where the test zoid did not fit. Lag is expressed in terms of the number of frames intervening between last preview and test zoid.

120.27, MSE = 1039.34). Finally, the effect of the presence of a preview was dependent on the status of the test zoid, with the preview effect being larger when the test zoid fit relative to when it did not (indirect, 140 ms vs. 131 ms, $F_{(1,14)} = 10.87$, MSE = 31.12; direct, 148 ms vs. 133 ms, $F_{(1,13)} = 5.08$, MSE = 181.28).

Having established that a preview made a difference, we next look to see whether having more than one preview made a difference, and whether the preview(s) had any interacting effects with other aspects of the design. The preview-present data were analyzed using two 2 (zoid type: ⊞ , ⊟) × 3 (number of previews: 1, 2) × 2 (orientation of the preview relative to the test piece: same, different) × 3 (lag, in frames, between the last preview and the test zoid: 0, 1, 2) × 2 (status of the test zoid: does fit, does not fit) repeated measures ANOVAs, one for each of the test conditions. A first result was an effect of number of previews: participants were faster with two previews than with one (indirect, 542 ms vs. 562 ms, $F_{(1,14)} = 101.02$, MSE = 353.96; direct, 857 ms vs. 872 ms, $F_{(1,13)} = 15.75$, MSE = 1172.30). The
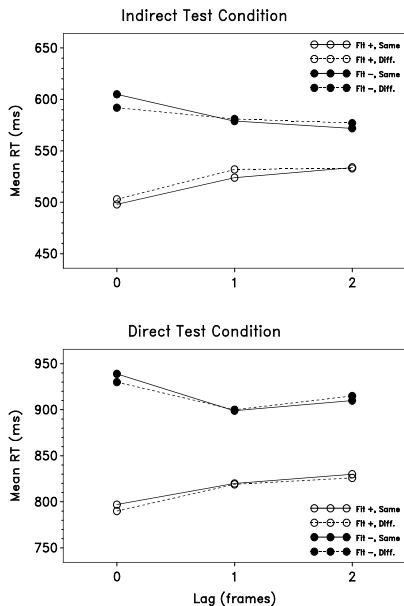
Figure 5: Preview-present trials: interaction of lag, fit status, and orientation of the test zoid.

Figure 6: Preview-present trials, direct test condition: interaction of lag, number of previews, and accuracy of the memory judgment.

lag between the last of the previews and the status of the test zoid interacted in both test conditions (see Figure 4): decreases in lag produced faster RTs when the test zoid fit (indirect, $F_{(2,28)} = 73.19$, MSE $= 379.39$; direct, $F_{(2,26)} = 21.72$, MSE $= 1488.90$) but produced longer RTs when the test zoid did not fit. Finally, there was an interaction among lag, status of the test zoid, and orientation of the preview relative to the test zoid (see Figure 5), though this interaction was reliable only for the indirect condition ($F_{(2,21)} = 4.21$, MSE $= 339.60$).

We next examined trials on which there were two previews. Half of these involved previews in one orientation and half involved previews in two orientations. Analysis revealed that seeing two orientations led to faster responses than seeing one orientation in both indirect (616 ms vs. 686 ms, $t_{(28)} = 4.23$) and direct (932 ms vs. 980 ms, $t_{(26)} = 4.18$) conditions.

If epistemic actions serve the mnemonic purposes we have suggested, then there might be a certain awareness on the part of the player as to mnemonic state while playing, which in turn suggests that the ability to monitor memory state may modulate the benefits of epistemic action. To assess this possibility, we examined the preview-present trials of the direct test condition, treating accuracy of memory
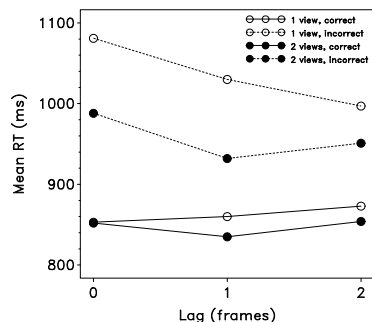
judgment as a random effect, using a 2 (memory accuracy: correct, incorrect) × 2 (zoid type: ⊞ , ⊟) × 3 (lag between last preview and test zoid: 0, 1, 2) × 2 (number of previews: 1, 2) repeated measures ANOVA. This analysis revealed the expected benefit of increasing the number of previews (899 ms for two vs. 946 ms for one, $F_{(1,24)} = 6.66$, MSE $= 1194.80$). It also revealed faster responses when participants accurately remembered the preview relative to when they did not consciously recall seeing a preview (855 ms vs. 997 ms, $F_{(1,24)} = 4.79$, MSE $= 96227.58$). Finally, there was an interaction among lag, number of previews, and the accuracy of the memory judgment (see Figure 6; $F_{(2,48)} = 4.02$, MSE $= 1851.61$).

Discussion

Our results show that when the test zoid and board are not oriented properly with respect to one another, one preview of the zoid in any orientation with respect to the board leads to faster responses than does no previews, and two previews lead to faster responses than does one preview. This was found for both indirect Tetris-like fit/no-fit judgment task as well as for direct memory recognition task. In addition, two previews with two different orientations produced faster responses than did two previews with the same orientation. Thus, under certain conditions, several orientations can prime two-dimensional shape recognition and use better than a single orientation can.

Note that response time was speeded up by a single preview in any of the three orientations relative to the test zoid and board. The benefit was not restricted to a preview 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). However, we have gone a step further than these by showing that *priming with several orientations is more effective than priming with a single orientation* under certain conditions.

The data also revealed that the benefit of previews, on trials in which the test zoid did fit, diminished as time elapsed between last preview and time of decision. This attenuation of the positive effects of previews suggests that the benefit of rotation for self-cueing may be restricted to a small window of time just prior to the final decision, which would be consistent with the reasonably rapid pace at which the game proceeds for skilled players. In contrast, on trials in which the test zoid did not fit, temporal proximity between last preview and judgment appeared to extract a cost, suggesting a strong specificity of the effect of previews to particular conditions of the game. Moreover, the data from the direct test condition revealed that accurate, conscious memory of a preview produced a benefit in responding, suggesting that players may have the ability to monitor their mnemonic state—as well as the state of the game—as play unfolds. One provocative idea is that epistemic actions may occur in response to players' assessment of a need for additional cueing.

Returning to the specific idea epistemic action in Tetris—the board-match function of rotation in particular—these results suggest that by rotating 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 & Kirsh, 1996). The data reported here show that several previews of the falling zoid sometimes 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. It also remains to be seen whether the time-cost of making an extra move is more than compensated by the benefit in RT. We are exploring both questions.

References

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

Cooper, L. A., Schacter, D. L., Ballesteros, S., & Moore, C. (1992). Priming and recognition of transformed three-dimensional objects: Effects of size and reflection. *Journal of Experimental Psychology: Learning, Memory & 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 & Cognition, 19*.

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. LEA.

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. LEA.

Maglio, P. P. & Wenger, M. J. (2000). Two views are better than one: Epistemic actions may prime. In *Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society*, Mahwah, NJ. LEA.

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 & Cognition, 21*, 1019–1036.

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

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