 Strategies in Visuospatial Working Memory for Learning Virtual Shapes

GLENN GORDON SMITH1*, ALBERT DIETER RITZHAUPT1 and EDWIN TJOE2

1Department of Secondary Education, University of South Florida, USA
2Technology & Society, Stony Brook University, USA

SUMMARY
This study investigated visuospatial working memory (WM) strategies people use to remember unfamiliar randomly generated shapes in the context of an interactive computer-based visuospatial WM task. In a three-phase experiment with random shapes, participants (n = 94) first interactively determined if two equivalent shapes were rotated or reflected; second, memorized the shape; and third, determined if an imprint in a profile view of the ground was a rotated, reflected imprint of the shape, or an imprint not matching the original shape. Participants self-reported these strategies: Key feature, shape interaction, association/elaboration, holistic/perspective, divide and conquer, mental rotation/reflection and others. Participants reporting key features strategy were significantly more accurate on the computer-based visuospatial WM task. These results highlight the importance of strategy in visuospatial WM. Copyright © 2009 John Wiley & Sons, Ltd.

INTRODUCTION
When people encounter new shapes in the context of interactive computer programs, how do they remember those shapes? What strategies do they use and how effective are they? With the proliferation of virtual worlds, including simulations in biology, chemistry, health sciences and computer gaming, people are likely to encounter more unfamiliar shapes for which they are often held accountable. It may be of interest, in the fields of psychology and education, to investigate which visuospatial working memory (WM) strategies people use to remember and work with novel shapes. In the current study, we investigated participants’ visuospatial short-term memory/working memory (STM/WM) strategies with shapes and geometry unfamiliar to the participants for two reasons. First, the question of what strategies people use to encode unfamiliar shapes is important educationally and vocationally. Second, the question has theoretical value. The authors hypothesize that the range of strategies, and their relative effectiveness, is broader than what is illustrated in the current research literature. Different geometries may engender different strategies, with different relative effectiveness.

*Correspondence to: Glenn Gordon Smith, Department of Secondary Education, University of South Florida, 4202 E. Fowler Ave., EDU 162, Tampa 33620-5650, USA. E-mail: gsmith@coedu.usf.edu

Copyright © 2009 John Wiley & Sons, Ltd.
Many studies investigating spatial strategies, such as Schultz (1991) and Just and Carpenter (1985) and Carpenter (1986), have used standardized tests of spatial visualization, where the spatial arrays are always visually present while participants mentally work with shapes. Thus, participants in these studies may actually avoid using visuospatial STM/WM altogether. The current study, investigating visuospatial STM/WM strategies, was designed to tap into visuospatial STM/WM by using a task where participants had to memorize shapes, and then make decisions about shapes, in the absence of pictures of the original shapes. The current study builds on a prior study investigating visuospatial STM/WM strategies which used a task involving common basic geometric shapes (Kyllonen, Lohman, & Woltz, 1984), but diverges from that study by using unfamiliar, random shapes. Thus, the current study also investigated the hypothesis that choice and effectiveness of visuospatial STM/WM strategies relates to the geometry of the shapes, and also the relative familiarity or unfamiliarity of the shapes.

From a theoretical standpoint, STM and WM may not be separable. Indeed for both verbal and visuospatial modalities, recent evidence (Unsworth & Engle, 2007) suggests that ‘the notion that STM and WM are largely different constructs is unwarranted’ (p. 1056). Further, there is a strong involvement of executive control in both visuospatial STM and visuospatial WM (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). Such involvement of executive control suggests alternatives in how shapes are remembered, and beckons more research on visuospatial STM/WM strategies. From a point of view of educators, it may be useful to instruct students on effective strategies for working with novel shapes.

Throughout our lifespan, first as infants, then young children and finally adults, we encounter new shapes. Every shape is new at some point. We are born with an empty filing cabinet that we gradually fill up with shapes. The ability to integrate new information or build new information structures is a critical part of learning. However, most new shapes are easily recognized variations or compositions of old shapes. The authors suggest and assume that occasionally, the cumulative differences between new shapes and old shapes are so great, that we may perceive a shape, not as a variation of old shapes, but as an unfamiliar and novel new shape.

With growing importance of life-long learning and the increased synthesis of disciplines, professionals may have to work with such perceived unfamiliar shapes. For example, a technician or a software engineer or instructional designer for a small start-up biotechnology company, with little chemistry knowledge, might have to identify proteins of the Human Immunodeficiency Virus (HIV). In the recreational realm, online computer game players may encounter unfamiliar shapes, for example, whimsical keys to open fantasy doors to imaginary realms. College students in chemistry classes often encounter complex shapes of molecules. Three-dimensional interactive visualization software is common in fields such as chemistry, geology and health-care, and interaction with virtual shapes is often involved in these virtual environments.

Research indicates that educators should present new content in the context of existing student knowledge, in terms of previous course material and students’ informal knowledge (Allen & Boykin, 1992; Au & Jordan, 1981; Boykin & Tom, 1985; Erickson & Mohatt, 1982; Shin, Schallert, & Savenye, 1994). This is consistent with the connectionist theory of memory (Craik & Lockhart, 1972). Going even farther back in the psychology literature, people learn primarily by assimilating new events into their existing schemas, or modifying existing schemas to accommodate new knowledge (Piaget & Inhelder, 1969). Further, adding a social dimension to connectionism creates constructivism, a currently prevalent educational philosophy, where students, mediated by teachers, collaboratively construct
their own new knowledge building on their existing knowledge structures (Cobb, 1994; Piaget & Inhelder, 1969; Vygotsky, 1962, 1978).

But what if the educational situation is not consistent with theoretical conditions? If the learning situation consists of unfamiliar shapes with little connection to prior knowledge structures, what strategies will people choose? How effective will these strategies be? From a practical point of view, workers, computer game players, and students may move into a new area with its own unique geometry. In the work arena, without the time or resources to truly educate workers in the discipline, companies may need to train technicians for a job involving a complex, unfamiliar geometry. Similar situations arise in recreational and academic arenas.

LITERATURE REVIEW

Visuospatial memory

STM is a limited capacity (seven plus or minus two pieces of information), temporary memory for holding information currently in the focus of a person’s attention (Miller, 1956). WM, on the other hand, concurrently stores as well as processes information held in STM for some cognitive task (Baddeley, 1974). WM can be divided functionally into four components, the central executive (conscious decision-making) and three slave systems, verbal, episodic (multi-modal) and visuospatial WM (or the ‘visuospatial sketchpad’) (Baddeley, 1986; Baddeley & Hitch, 1999; Baddeley, 2000). The central executive directs and coordinates verbal and visuospatial memory, and the episodic buffer.

Although more recent opinions suggest otherwise (Unsworth & Engle, 2007), in the verbal WM system, there is some evidence that a functional distinction can be made between STM and WM (Baddeley, 1986; Baddeley & Hitch, 1999). Across modalities, the differentiation of STM and WM is tenuous (Unsworth and Engle, 2007), but in visuospatial WM, the distinction between the visuospatial STM, WM and the central executive is even murkier than in the verbal modality. In factor analytic studies and interference studies, visuospatial STM and visuospatial WM (comprising visuospatial STM and related executive functioning/attention) are practically indistinguishable (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; Smyth & Pelky, 1992). Thus, one may speak of a visuospatial WM cluster which includes executive functioning. It requires a decision, and effort of will to conjure up visual mental images (Hasher & Zacks, 1979), and then ongoing attention to maintain them in the mind’s eye (Awh & Jonides, 2001; Shah & Mijake, 1996). Consequently, executive involvement is required for visual mental imagery.

This involvement of executive control in visuospatial WM also suggests the use of strategy. If conscious decision-making is required to maintain temporary storage of shapes, then there must be more than one method for temporarily storing shapes. Further, there are choices of different operations that can be performed on shapes held in visuospatial WM, including transformations like mental rotation and changes in perspective discussed in the spatial abilities literature.

Visuospatial cognitive tasks

In terms of human spatial skills, there are three basic reference systems for visuospatial images: (a) eye, (b) effector (hand and foot) and (c) object; and three corresponding spatial
transformations: (a) perspective, (b) effector-based and (c) object-based (Zacks & Michelon, 2005). These are related to a number of corresponding human spatial abilities. Table 1 shows an overview. If one takes the evolutionary view that cognition has evolved for humans to simulate action in advance of performing such action (Glenberg et al., 2007), then human spatial abilities and their reference systems correspond to human bodily affordances: (a) traveling through landscapes and viewing others doing so, (b) large scale body movements with arms and legs (c) hand manipulation of small objects afforded by opposable thumbs.

In perspective transformations, the human eye is the centre of the frame of reference. A perspective transformation means changing the imagined location of the human eye in viewing a landscape, e.g. a photographer standing under the old clock tower in a city square imagining how a photograph might look as viewed from the opposite end of the square, looking towards the clock tower and looking towards the photographer’s position. Perspective transformations correspond to one of the basic human spatial abilities, spatial orientation, which is defined as the ability to visualize how a scene looks from a different point of view (Carpenter & Just, 1986; Lohman, 1988; Pellegrino & Kail, 1982). Spatial orientation involves two essential components: (a) a change in the orientation or position of the viewer and (b) a spatial array such as a landscape, on a larger scale than the human being, such that the viewer might be a part of the environment.

In effector-based transformations, frame of reference is focussed on a foot, hand or other body effector. These effector-based transformations correspond to kinesthetic or motor mental imagery. Generally, effector-based transformations relate to an individual imagining performing a body movement, such as swinging a golf club.

In object-based transformations, the frame of reference is centred on a small object of hand-manipulable scale, and that object is transformed relative to that object’s frame of reference. An example is imagining a toy car spun around, or flipped over. Object-based transformations correspond to two other basic spatial abilities—mental rotation and spatial visualization. Mental rotation (MR) is the ability to imagine rotating one shape into alignment with another (Shepard & Cooper, 1982). Another basic spatial ability (also relating to object-based transformations) is spatial visualization, the ability to solve multi-step problems involving complex shapes, or configurations of shapes (Linn & Petersen, 1985; Smith, 1998; Zimowski & Wothke, 1986).

There is one more human spatial skill, spatial perception, which has no corresponding spatial transformation or reference system, but which figures in visuospatial WM strategy. Spatial perception is a ability to discern spatial relationships in the presence of distracting information (Linn and Petersen, 1985).

The current study focusses on visuospatial WM in the context of object-centred transformations and their associated spatial abilities, mental rotation and spatial visualization.

<table>
<thead>
<tr>
<th>Centre of frame of reference</th>
<th>Transformation</th>
<th>Spatial abilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye</td>
<td>Perspective</td>
<td>Spatial orientation</td>
</tr>
<tr>
<td>Hand or foot</td>
<td>Effector-based</td>
<td>Kinesthetic or motor mental imagery</td>
</tr>
<tr>
<td>Small object (potentially hand-held)</td>
<td>Object-based</td>
<td>Mental rotation, spatial relations, spatial visualization</td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td>Spatial perception</td>
</tr>
</tbody>
</table>
Visuospatial strategies

The spatial abilities research literature reports a limited number of strategies used on psychometric tests of spatial abilities. See Table 2 for an overview of strategies. Schultz (1991) identified three strategies used on a number of standardized spatial tests: (1) mentally move object (analogous to imagining moving a hand-held object), (2) mentally move self (a person imagining moving relative to a larger environment) and (3) analyse in terms of key features (a logical process that involves verifying whether key features in one spatial array appear in the same relative position in another array).

Just and Carpenter (1985) suggested that three strategies were used in the cube comparisons test, a test requiring participants to indicate whether two drawings of cubes with letters on each face are of the same cube rotated in three-dimensional space, or different cubes. These strategies included a mental rotation strategy, a perspective-taking strategy, and a strategy comparing orientation-free descriptions. Burin, Delgado and Prieto (2000) found holistic versus analytic strategies for performance of a ‘formboard’ style test in which participants determine if a small number of polygons could hypothetically be assembled into a larger target or two-dimensional shape. Holistic strategies involve the mental transformation of a whole shape, for example, visualizing in the mind’s eye that the Washington Monument rotates 90 degrees. Analytic strategies do not involve visualizing shapes or transformations, but rather involve logical deduction based on properties of shapes, for example, reasoning that a closed shape composed of four equal line segments and four right angles is a square. See Table 2 for an overview of these strategies and their relationship to spatial abilities.

According to Hegarty (2009), one hallmark of spatial expertise is flexibility of strategies. Visualization of shapes and other holistic strategies are cognitively demanding. Thus, while a spatially skilled person needs to be able to visualize and mentally transform shapes when necessary, often an analytical or abstract strategy works just as well with less effort. So for example a well-known mechanical reasoning task involves viewing a diagram of a series of gears and determining if you turn the first gear clockwise, what direction does the fifth gear in line turn (Hegarty, 1992). A spatially skilled person initially solves this class of problem by mentally rotating the gears. However, after solving a few such problems, it becomes apparent that all the odd gears will rotate in one direction, the even gears in the opposite direction. Visualization of shapes may be necessary to discover this rule, but application of this rule will save cognitive work. The same principle holds true in disciplines such as chemistry where students may expend efforts on visualization, but experts typically employ analytical strategies (Stieff, 2009). Even psychometric tests designed to measure holistic visualization ability often lend themselves to analytic strategies. Thirty-eight per cent of people solve items on the Vandenberg and Kuse Mental Rotation test, not by mental rotation, but by counting cubes or comparing key features on shapes (Hegarty, 2009).

Table 2. An overview of spatial strategies from the spatial research literature

<table>
<thead>
<tr>
<th>Type</th>
<th>Strategies</th>
<th>Spatial skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holistic</td>
<td>Mentally move object, Mental Rotation</td>
<td>Mental rotation, Spatial visualization</td>
</tr>
<tr>
<td></td>
<td>Mentally move self, Perspective taking</td>
<td>Spatial orientation</td>
</tr>
<tr>
<td>Analytical</td>
<td>Analyse in terms of key features, Decomposition</td>
<td>Spatial perception</td>
</tr>
<tr>
<td></td>
<td>Comparing orientation-free description</td>
<td></td>
</tr>
</tbody>
</table>
As noted in the introduction, many studies of spatial strategies keep the spatial stimulus in full visual view, such that the person can refer to it. Thus, the person does not necessarily have to fully encode the shape in WM. One notable exception is Kyllonen, Lohman, and Woltz (1984), who investigated spatial strategies for a task that involved visuospatial WM remembering (encoding) of shapes. It is described in some detail as the current study reacts to and builds on it. Kyllonen et al.’s (1984) task (see Figure 1) involved three phases: (1) WM memorization or encoding of a shape (referred to as figure A), (2) synthesis of the now absent figure A with two new figures and (3) comparison of a test figure with the absent synthesized shape. Kyllonen et al. (1984) investigated strategies through componential analysis. They proposed certain strategy models, calculated how long the use of such a strategy would take for specific shapes used, and then based on elapsed times for encoding, synthesis and comparison phases, used multiple regression to deduce which strategies were actually used. They also evaluated different strategies in terms of performance, including speed of encoding and accuracy on the comparison task. The strategy models for visuospatial WM encoding that they investigated included: (1) key feature, (2) decomposition into sub-shapes, and (3) verbal labelling. The decomposition encoding strategy was optimal for speed of encoding and accuracy on the comparison task. While acknowledging the importance of this seminal study, the current authors offer the following critiques. First, by using a componential hypothesis-testing approach, instead of a more qualitative approach, such as self-reporting of strategies, some strategies may have been overlooked. More importantly, the shapes used for encoding lend themselves to a decomposition strategy. A close examination of the shape from Kyllonen et al. (1984) to be encoded (see Figure 1) reveals that it lends itself to decomposition into a triangle and two rectangles. This may be a byproduct of the premeditated construction of these shapes. Figures A, B and C (of Figure 1) are the result of backwardly decomposing the final test probe (a simple right triangle) into slightly more complex shapes. The current authors hypothesize that the relative speed, accuracy, as well as frequency, of encoding strategies relates to the geometry of the shapes to be encoded. This is especially so if, as in Kyllonen et al. (1984), the shapes involved are constructed from familiar, simple shapes, such as squares, rectangles, equilateral or isosceles triangles. Shapes created by computers or from some natural processes may be based on less familiar geometric patterns, and therefore, they may lend themselves to different strategies. The geometry of the shapes is an affordance (Gibson, 1979) for visuospatial WM strategies.

**Gaps in research on visuospatial STM/WM strategies**

One problem with the previous research is the indirect and deductive methods used in the investigation of visuospatial WM. Methods such as factor analysis, componential analysis and brain imaging techniques, while effective in establishing broad patterns, are less
effective for establishing specific strategies. There is a need for more direct research methods that uncover which strategies are used in specific visuospatial WM tasks.

Secondly, the relationship between the geometry of the shapes and the strategies used, and their relative effectiveness, remains unexplored. The visuospatial strategies explored do not explain what people use when confronted with unfamiliar or seemingly random geometries.

Further, the use of standardized tests of spatial abilities to investigate spatial strategies may not be ecologically valid. It is important to investigate strategies for visuospatial WM in educational, academic or work activities. Many of the visuospatial WM situations occur in the real world, navigation/way-finding, imagined or actual manipulation of hand-held or larger objects. However, increasingly visuospatial WM situations occur in the virtual work through interactive computer graphics animations and simulations. Such spatial situations in the virtual world may not currently be considered naturalistic, but if current work and education trends continue, interacting with computers will eventually be considered as naturalistic as reading books. Interaction, the manipulation of virtual shapes via computer input devices, needs to be a part of study materials and settings. The involvement of interaction may have some unexpected effects on visuospatial WM strategy. Thus, it is important to investigate visuospatial WM strategies in contexts involving interaction with virtual shapes.

Research questions

The current study assumes the existence of strategies in visuospatial WM tasks. The authors are interested in what strategies people use for visuospatial WM of unfamiliar shapes in common computer-based tasks that also involve interaction and spatial abilities such as mental rotation and spatial visualization (Juhel, 1991). When presented with common computer-based visuospatial WM tasks, will people try to remember shapes by connecting them to previous familiar shapes, or will they use strategies that are based less explicitly on familiar shapes? How do different strategies compare in terms of performance, such as accuracy and speed?

Thus, the main research questions were: (a) what strategies do people use for common computer-based visuospatial WM tasks involving unfamiliar shapes, and (b) how effective are these strategies in terms of task accuracy and speed?

Based on an emphasis in education on building on prior knowledge, and emphasis on verbal and semantic knowledge over visuospatial skills (Smith, 1964), one might expect people to predominantly use simple knowledge-based strategies. For example, people might suggest that a shape looks like a dog’s head. However, the current investigators hypothesize that such strategies might not be effective as they gloss over finer details of shapes. Greatest retention of detail and thus best performance might come with analog or holistic strategies (Zimowski & Wothke, 1986; Burin, Delgado, & Prieto, 2000) involving memorization of whole shapes. People might also employ key feature strategies, memorizing parts of shapes, as these strategies might be less cognitively demanding. However, participants’ judgment of which features are really important might not be dependable. Therefore, the investigators hypothesize that feature strategies might be less effective than holistic/analog strategies.

The investigators hypothesize that some people will employ strategies related to previous experience and semantic knowledge and categories; while others will employ strategies using relatively less experienced-based semantic categories. The investigators...
also hypothesize that the geometry of the shapes influences choice of strategy and relative effectiveness of strategy. Thus, using different shapes than used by Kyllonen et al. (1984), will produce different strategies and different relative effectiveness of strategies. The geometry of a shape is an affordance (Gibson, 1979) that may lend itself better to one strategy or another. People will choose a strategy that is easier or more effective for that particular shape or geometry.

METHOD

Participants

Ninety-four undergraduate students from a large public university in the northeast United States participated in the study, receiving a small portion of extra credit in their courses for their participation. The participants ranged in age from 18 to 39 ($M = 20.1$, $SD = 2.95$). Sixty-nine per cent of the participants were male. One-hundred per cent reported owning a personal computer, and their average frequency of computer game play was 2.23 sessions per week ($SD = 2.21$). Eighty-three per cent of the participants were right-handed.

Materials and tasks

For this study, the investigators wanted to investigate visuospatial WM strategies that might occur in a variety of interactive computer environments, but especially those involving chemistry and crystallography visualization. These programs, such as ChemDraw (ChemDraw, 2009), Chime (Chime Pro, 2009), Rasmol (Bernstein, 2005), MolviZ (Martz, 2005), Jmol (2008) (Martz, 2006), FirstGlance in Jmol (2008) (Martz, 2006), ConSurf (Landau, Mayrose, Rosenberg, Glaser, Martz, Pupko, & Ben-Tal, 2006), Polyview-3D (Porollo & Meller, 2006) and Mercury for interactive visualization of crystallography (Cambridge Crystallographic Data Centre, 2004) involve users in multiple steps with interaction with shapes, spatial cognition such as mental rotation and spatial perception, and visuospatial WM of new shapes. For this study, the investigators wanted to look at strategies employed in a computer-based task, involving interaction, similar to these types of chemistry visualization programs, and other programs and games in which people interact with shapes. The investigators wanted a task involving interaction and cognitive operations similar to those in chemistry visualization programs and a task similar to identifying right and left hand versions of molecules in chemistry. In designing such a task, the investigators determined that participants should: (a) interact with unfamiliar shapes, (b) encode the unfamiliar shapes in STM and subsequently use them in a complex accountable spatial task (visuospatial WM), and (c) in the accountable spatial task, recognize the shape or transformed parts of the shape in another context and differentiate the shape from other shapes.

The investigators programmed a Java 2 Applet that generated random polygons of six vertices and presented them to participants in three phases shown in Figures 2–4. The use of randomly generated polygons was used to avoid commonly seen and familiar basic shapes. Within each trial, the same shape was used in each phase. Between different trials, different shapes were used. Thus, each trial had its own unique shape that was used (explicitly or implicitly) throughout each of the three phases.
Phase 1: Participants were presented with two shapes that were rotated or reflected versions of the same shape. Participants had to determine, by interactively rotating one shape with the mouse, whether the second shape was a rotated-only version of the first shape, or reflected in addition to being rotated (see Figure 2).

The participant indicated whether the shape was rotated or reflected by clicking either on a button labelled ‘turn’ or a button labelled ‘flip.’ The buttons for ‘flip’ or ‘turn’ were actually grayed and unavailable until the participant interactively rotated one of the two shapes.
shapes with the mouse. Participants were informed as to the correctness of their response with visual animated feedback, such as showing the second shape rotating onto the first, or the second shape rotating a full 360 degrees without coinciding with the first shape. As determined by a randomizing function, the odds of the second shape being flipped (reflected) were 50-50. Interaction was included in phase 1 to make it similar to interactive chemistry modelling software involving extensive interaction to transform shapes. Also, often in computer games and recreational virtual worlds, interaction involves the transformation of shapes.

**Phase 2:** Participants were presented with the same shape as the one on the left in phase 1 and asked to remember the shape. Participants were notified that they could interactively rotate the shape with the mouse, if that helped them to remember the shape. Again, interaction was included in phase 2, to be consistent with interactive programs that involve interaction with shapes.

**Phase 3:** Phase 3 either explicitly or implicitly involved the shape from phases 1 and 2. Participants were presented with a silhouette showing what appeared to be a profile of a hole in the ground and had to decide if the hole was an imprint of the shape from phases 1 and 2 that was: (a) rotated only, (b) reflected as well as rotated or (c) the imprint of some other shape. Participants indicated their decision by clicking on one of the three buttons, labelled ‘turned’, ‘flipped’ or ‘completely different.’ As in phase 1, in phase 3 the participant was given animated feedback about the correctness of their answer.

Whether the ‘hole in the ground’ was a turned, flipped version of the shape from phase 1 and 2, or completely different was determined by a randomizing function. Within phase 3 in a given trial, the odds of each of the three possibilities (turned, flipped, or completely different) were equal (e.g. 1 in 3).

Phase 3 is similar to enzyme kinematics: D-glucose *versus* L-glucose and how these do or do not fit into the substrate. Phase 3 is, by definition, a visuospatial WM task. As mentioned earlier, WM is defined as involving the concurrent storage as well as processing of information held in STM for some cognitive task (Baddeley and Hitch, 1999). To perform the task in phase 3, the participant must hold in visuospatial STM the shape from phases 1 and 2, and concurrently make a judgment about the shape of the imprinted hole. The participant concurrently stores visuospatial information as well as processes visuospatial information held in STM for some visuospatial cognitive task.

**Procedures**

Since the aim was to investigate what strategies were used for visuospatial WM tasks, the study used verbal reports. Since concurrent verbal reports (thinking aloud during tasks) can change the nature of non-verbal processing (e.g. translation from non-verbal to verbal modality) (Ericsson & Simon, 1993), retroactive verbal reports were used. Verbal overshadowing of visuospatial encoding can interfere with long-term memory (LTM) access of visuospatial representations. Brandimonte et al. (1992) suggested that phonological recording in verbal STM during learning prompts the establishment of some form of verbal or propositional code in LTM which is detrimental when the task to be performed requires the recovery of visual information (p. 455). In Dunlosky and Kane’s (2007) study, they used direct participant reporting of strategies, including participants reporting after each set of items, and fully retrospective participant reporting of strategies after completion of the span test. Fully retrospective reports were just as reliable as the by-
set reports and did not suffer from the potential disadvantage of affecting subsequent performance. The current study also used retrospective reports.

All participants were given a scripted introduction to the phases, and notified of a final survey and an exit interview following the session. Secondly, all participants completed a short computer-facilitated background survey collecting relevant information (e.g. frequency of game play, gender, age). Next, participants finished 30 trials, each comprised of phases 1, 2 and 3. Next, they completed a two question computer-facilitated survey on strategies to complete phases 2 and 3, and then completed a brief oral exit interview with an investigator. Visual cues (screenshots) of the phases were used to elicit the information. The purpose of the exit interview was to probe for more information related to the strategies they employed in each phase; and to triangulate the descriptions with the survey responses.

Reliability of measures

Classical methods of reliability (e.g. Cronbach’s Alpha) were not used to estimate the reliability of phase 1 and phase 3 tasks because the shapes were randomly generated as oppose to each participant receiving the same shapes. A suitable alternative approach to estimate reliability is generalizability theory, or more specifically, a G-study (Crocker & Algina, 1986). This G-study used the trials as a single facet and participants as the object of measurement; thus, the design uses the trials, participants and their interactions to estimate the generalizability coefficient, which is analogous to a reliability coefficient (Crocker & Algina, 1986). Specifically, there were 94 participants and 30 randomly generated trials in both phase 1 and phase 3.

In terms of phase 1, approximately 85 per cent of the variability is explained by the interaction of the object of measure, the participant and the one facet, trial. This is an indication that increasing the number of trials increases the generalizability coefficient. The generalizability coefficient calculated under the assumption of a single item was calculated at $\rho^2_{r} = 0.14$. When accounting for 30 randomly assigned items, the generalizability coefficient increases substantially, $\rho^2_{r} = .83$. Thus, the phase 1 task was a reliable measure for these data.

Phase 3 resulted in 95 per cent of the variability explained by the interaction of participants and trials. Again, this provides strong evidence that increasing the number of randomly generated trials positively impacts the generalizability coefficient. The generalizability coefficient calculated under the assumption of a single item, was calculated at $\rho^2_{r} = 0.04$. The generalizability coefficient increased to, $\rho^2_{r} = .57$ when accounting for 30 randomly generated trials. The phase 3 task may have benefited from an increased number of trials for these data. Specifically for these data, 51 random trials in phase 3 would have resulted in a more desirable generalizability coefficient (>=0.7).

Qualitative analysis

The investigators analysed the self-reported strategies and exit interviews used in phase 2 (memorizing shape) and phase 3 (visuospatial WM task) with an inductive latent content analysis approach (Tashakkori & Teddlie, 1998). Latent content analysis was used to identify, code and categorize the primary strategies in the data. The investigators first sought the meaning of the survey responses and exit interviews within the context of all the data (Mayan, 2001), developed a categorization scheme, and then coded the data according to the categories using this scheme (Polit & Beck, 2005). Two investigators independently
coded the strategies for all the data. Cohen’s Kappa was calculated for each of the strategies to control for chance. Landis and Koch (1977) provide guidelines for evaluating Cohen’s Kappa: 0.21-0.40, ‘Fair’; 0.41-0.60, ‘Moderate’; 0.61-0.80, ‘Substantial’; 0.81-1.00 ‘Almost perfect.’ Using these guidelines, the first occasion Cohen’s Kappa indicate that one of the strategies was fair, one was moderate, eight were substantial and two were almost perfect (See Tables 3 and 4). On the second occasion, the investigators recoded the data in light of the Cohen’s Kappas until inter-rater agreement was equal to 100 per cent and a final set of strategies was devised.

RESULTS

Phase 1
On average, participants completed phase 1 with a mean accuracy of 90 per cent ($SD = .12$) and with a mean latency of 6.9 seconds ($SD = 2.60$) per trial. On phase 1 (an interactive rotation task), males were significantly more accurate ($M = .92$, $SD = .09$), than were females ($M = .86$, $SD = .17$), $t(92) = −2.12$, $p < .01$ (two-tailed), $d = 0.16$. This was the only significant gender effect in the data.

Phase 2 (Encode)
In phase 2 (memorizing the shape), the average time participants took to memorize a shape in phase 2 was 6.2 seconds ($SD = 5.42$). The average number of strategies used in phase 2 was 1.27 ($SD = 0.63$). The following results address the main research question, ‘what strategies do people use for common computer-based visuospatial STM and WM (visuospatial sketchpad) tasks involving unfamiliar shapes’? The strategies are listed in order of frequency with a definition, typical quote and the Cohen’s Kappa from the first occasion of coding in Table 3.

Phase 3 (Decode)
In phase 3 (typical computer-based visuospatial WM task), participants were asked to determine whether the ‘hole’ in the silhouette matched the shape from phase 2 in: (a) rotated form, (b) reflected form or (c) a completely different shape. For the phase 3 task, the mean accuracy was 61 per cent ($SD = 0.13$), while the mean latency was 3.73 seconds ($SD = 2.43$). Because there were three possible choices in phase 3, a mean accuracy of 61 per cent is well above random guessing, which would be 33 per cent.

The average number of strategies reported used in phase 3 was 1.1 ($SD = 0.64$). The strategies employed in phase 3, as determined by qualitative categorization of self-reported strategies from survey items and exit interviews, are shown in Table 4. With the exception of process of elimination, all strategies reported in phase 3 were also reported in phase 2. Table 4 shows the strategies in order of frequency with a definition, typical quote and the Cohen’s Kappa from the initial coding.

Analysis of strategies on performance
The second research question was how effective were the various strategies on the visuospatial WM task. To answer this question, the investigators conducted a series of
<table>
<thead>
<tr>
<th>Strategy/Kappa</th>
<th>Definition</th>
<th>Typical quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key feature $\kappa = .77$</td>
<td>Participants memorized distinctive features of the shape.</td>
<td>‘I chose a part of the shape that stands out.’</td>
</tr>
<tr>
<td>Shape interaction $\kappa = .67$</td>
<td>Participants used the mouse to interactively rotate the shape while memorizing it.</td>
<td>‘I rotated the shape bit by bit and tried to picture what it would look like on the following page.’</td>
</tr>
<tr>
<td>Association $\kappa = .96$</td>
<td>Participants elaborated on the shape by associating it with a familiar object.</td>
<td>‘I tried to relate the shape to a real object.’</td>
</tr>
<tr>
<td>Holistic/Perspective $\kappa = .79$</td>
<td>The participant observed the image holistically as oppose to breaking it apart or focusing on only a key feature. The participants would usually describe looking at the entire image from different perspectives.</td>
<td>‘I spun it and tried to remember the orientation of the angles when viewed in a different position.’</td>
</tr>
<tr>
<td>Divide and conquer $\kappa = .73$</td>
<td>Participants mentally divided the shape into smaller pieces to memorize it.</td>
<td>‘Sometimes I divide the shape into several shapes.’</td>
</tr>
<tr>
<td>Bottom surface $\kappa = 1$</td>
<td>Participants interactively rotated the longest side of the shape to the bottom and aligned it horizontally. The investigators considered this related to key feature, since the longest edge was a key feature. However, the interactive rotation in a specific orientation distinguished it.</td>
<td>‘I just put the longest line as a bottom line and then try to figure out which way it is skewed.’</td>
</tr>
<tr>
<td>Rotation and reflection $\kappa = .37$</td>
<td>In order to remember the shape, participants mentally rotated or reflected the shape.</td>
<td>‘I memorize the image and try to picture it rotated and flipped.’</td>
</tr>
<tr>
<td>Outlining $\kappa = .67$</td>
<td>Drawing with their fingers on the screen or the table, participants outlined the shape.</td>
<td>‘With my finger, I tried to draw the shape three times.’</td>
</tr>
</tbody>
</table>
independent samples t-tests with the reported use of each strategy in phases 2 and 3 as the independent (or grouping) variable and the accuracy and latency of the phase 3 tasks as the dependent variables. In each t-test, the binary independent variable was the presence or absence of a reported strategy. The dependent variables were accuracy and speed on the visuospatial WM task (phase 3). See Tables 5 and 6 for an overview. Participants reporting the use of the key features strategy were significantly more accurate in the phase 3 task. Specifically, participants reporting use of key features strategy in phase 2 (memorization) were significantly more accurate on the phase 3 visuospatial WM task ($M = 0.65, SD = 0.14$), than those who did not ($M = 0.58, SD = 0.12$), $t(79) = 2.41, p = 0.02$ (two-tailed),

Table 4. Phase 3 (decoding) strategy, definition, and example quotation

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Definition</th>
<th>Typical quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key feature $\kappa = 0.67$</td>
<td>Participants memorized distinctive features of the shape.</td>
<td>‘I try to fit the points into the figure and if it fit in the same way it was shown it was turned, if it fit in the opposite way then it was flipped, and if the points didn’t match at all it was totally different.’</td>
</tr>
<tr>
<td>Association $\kappa = 0.71$</td>
<td>Participants elaborated on the shape by associating it with a familiar object.</td>
<td>‘If in the second task the picture had certain features or if it reminded you of something, you would be easily able to tell what it was and if it was flipped or not.’</td>
</tr>
<tr>
<td>Rotation and reflection $\kappa = 0.76$</td>
<td>In order to remember the shape, participants mentally rotated or reflected the shape.</td>
<td>‘I would rotate and flip the shape in my head to see if it could fit it into the hole.’</td>
</tr>
<tr>
<td>Process of elimination $\kappa = 0.58$</td>
<td>Participants described a systematic method for eliminating possible outcomes.</td>
<td>‘First I determine if the silhouette matches the shape of the previous figure. If so, then I know the answer must be ‘Turn’ or ‘Flip.’ If not, then it is a ‘Totally Different’ figure. If it does match, then I compare the different orientations the shape.’</td>
</tr>
</tbody>
</table>

Table 5. Number and percentage of participants reporting strategies, and latency for phase 2 ($n = 94$)

<table>
<thead>
<tr>
<th>Encoding Strategy</th>
<th>n $^*$</th>
<th>%</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key features</td>
<td>31</td>
<td>33</td>
<td>6.90</td>
<td>5.75</td>
</tr>
<tr>
<td>Shape interaction</td>
<td>30</td>
<td>32</td>
<td>7.61</td>
<td>6.13</td>
</tr>
<tr>
<td>Association</td>
<td>16</td>
<td>17</td>
<td>6.19</td>
<td>3.85</td>
</tr>
<tr>
<td>Holistic/Perspective</td>
<td>12</td>
<td>13</td>
<td>8.48</td>
<td>6.50</td>
</tr>
<tr>
<td>Divide &amp; conquer</td>
<td>10</td>
<td>11</td>
<td>5.90</td>
<td>4.29</td>
</tr>
<tr>
<td>Bottom surface</td>
<td>4</td>
<td>4</td>
<td>4.92</td>
<td>1.36</td>
</tr>
<tr>
<td>Rotation reflection</td>
<td>4</td>
<td>4</td>
<td>5.50</td>
<td>6.19</td>
</tr>
<tr>
<td>Outlining</td>
<td>1</td>
<td>1</td>
<td>7.08</td>
<td>—</td>
</tr>
</tbody>
</table>

$^*$It is possible that a participant reported more than one strategy, and consequently, is classified into all strategies identified. The mean and standard deviations reflect participants using more than one strategy.
Similar, participants reporting using key features strategy in phase 3 were also significantly more accurate in the phase 3 (M = .66, SD = .149), than those who did not (M = .59, SD = .12), t(92) = 2.35, p = .02 (two-tailed), d = 0.52. The reported use of the other strategies did not contribute significantly to accuracy or speed on the phase 3 visuospatial WM task.

To investigate contribution of other auxiliary variables on the performance of the visuospatial WM task (phase 3), Pearson product-moment correlation coefficients were computed. A correlation matrix is provided in Table 7. The number of reported strategies used by a participant for phase 2 (M = 1.27, SD = .63, n = 94) was significantly correlated with their accuracy on phase 3 (M = .61, SD = .13), r = .29, p = .005. Latency for phase 2 (M = 6.2, SD = 5.42) was also significantly correlated with phase 3 accuracy, r = .21, p = .047, which indicates that participants who spent more time memorizing in phase 2 were also more accurate in phase 3. In terms of demographic variables, the frequency of computer game play sessions per week (M = 2.23, SD = 2.2,) was significantly correlated with accuracy on the phase 3 task, r = .22, p = .03.

The strong correlation between the frequency of computer game play sessions per week and the accuracy of phase 3 task indicates that computer game playing experience is potentially a confounding variable. To test this hypothesis, the data were entered into an ANCOVA to partial out the effects of game play experience in an effort to test the more durable effects of the key features strategy. The data were retested specifically for the

Table 6. Number and percentage of participants reporting strategies, and latency and accuracy for phase 3 (n = 94)

<table>
<thead>
<tr>
<th>Decoding strategy</th>
<th>n*</th>
<th>%</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation reflection</td>
<td>34</td>
<td>36</td>
<td>0.62</td>
<td>0.12</td>
<td>4.34</td>
<td>3.55</td>
</tr>
<tr>
<td>Process of elimination</td>
<td>29</td>
<td>31</td>
<td>0.59</td>
<td>0.14</td>
<td>3.45</td>
<td>1.32</td>
</tr>
<tr>
<td>Key feature</td>
<td>28</td>
<td>30</td>
<td>0.66</td>
<td>0.15</td>
<td>3.46</td>
<td>0.86</td>
</tr>
<tr>
<td>Association</td>
<td>11</td>
<td>12</td>
<td>0.62</td>
<td>0.10</td>
<td>4.22</td>
<td>1.83</td>
</tr>
</tbody>
</table>

*It is possible that a participant reported more than one strategy, and consequently, is classified into all strategies identified. The mean and standard deviations reflect participants using more than one strategy.

Table 7. Correlation matrix of relevant auxiliary and demographic variables with performance variables (latency and accuracy)

<table>
<thead>
<tr>
<th>Pearson correlations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participant age</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Game play frequency</td>
<td>-0.22*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Phase 1 accuracy</td>
<td>-0.02</td>
<td>0.01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Phase 1 latency</td>
<td>0.15</td>
<td>-0.05</td>
<td>-0.06</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Phase 2 latency</td>
<td>0.22*</td>
<td>-0.13</td>
<td>0.25*</td>
<td>0.50**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Phase 3 accuracy</td>
<td>0.05</td>
<td>0.22*</td>
<td>0.40**</td>
<td>-0.03</td>
<td>0.21*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Phase 3 latency</td>
<td>0.22*</td>
<td>0.01</td>
<td>0.19</td>
<td>0.46**</td>
<td>0.73**</td>
<td>0.19</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. # strategies phase 1</td>
<td>-0.08</td>
<td>0.02</td>
<td>0.29**</td>
<td>-0.22*</td>
<td>0.18</td>
<td>0.29**</td>
<td>0.06</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9. # strategies phase 2</td>
<td>-0.05</td>
<td>-0.07</td>
<td>0.05</td>
<td>0.08</td>
<td>0.00</td>
<td>0.18</td>
<td>0.07</td>
<td>0.26*</td>
<td>1</td>
</tr>
</tbody>
</table>

*Indicates significant at 0.05 (two-tailed).
**Indicates significant at 0.01 (two-tailed).
participants that indicated a key features strategy for both phase 2 and phase 3. The results indicate that the key features strategies reported in phase 2 and phase 3 were still significantly related to accuracy in phase 3 while statistically controlling for the influence of computer game play sessions per week at $F(1, 91) = 4.49, p = .04$ and $F(1, 91) = 6.80, p = .01$, respectively. Thus, computer game play experience did not confound the relationship between the key features strategy and the accuracy on the phase 3 task.

DISCUSSION

One interesting result of the current study is the number of strategies reported seems greater than the small number reported in the spatial skills strategy literature. The current authors have noticed strategies such as association, bottom surface, interaction and outlining mentioned rarely, if at all, in other spatial skills strategy studies. The current study may have elicited more strategies because the task forced participants to remember shapes (or aspects of shapes) and then use that information in the absence of the original visual stimuli. Also, the unfamiliarity of the shapes may have forced participants to be more resourceful in remembering shapes. Further, the availability of interaction in the encoding phase provided potential for other strategies. This suggests that further ranges of experimental tasks might uncover still more strategies.

Never-the-less, one interesting result from the current study (which used random unfamiliar shapes) was that those who reported using the key features strategies in phase 2 (encoding) and phase 3 (decoding) were significantly more accurate in the phase 3 (visuospatial WM) task, even when controlling for game play experience. This is in contrast to the results reported by Kyllonen et al. (1984) who used basic and familiar shapes and found that decomposition strategies had the optimal speed-accuracy trade-off. This confirms the current investigators’ hypothesis that the geometry of shape and its relative familiarity influences relative effectiveness of strategies.

An association strategy, relating the random shape with a common object was the third most common strategy. This supports the hypothesis that some people would employ some semantic category-based strategies for remembering unfamiliar shapes in typical computer-based visuospatial WM tasks.

While participants self-reporting key features strategy were more accurate on the visuospatial WM task, those reporting association strategies enjoyed no such advantage. Associating the unfamiliar shape with shapes of more familiar objects did not improve either their speed or accuracy of visuospatial WM. The authors speculate that the association strategy over-simplifies the encoding. Features of the shape that resemble the known object are emphasized; features diverging from the known object are ignored. The known object is a schema; the shape is memorized in terms of features resembling invariant features of the schema. Those features not resembling the schema are forgotten. So much of education involves building on prior knowledge. Here is a case where connecting to prior knowledge may not be most effective. In this study, analysing the intrinsic properties of shapes worked better than other strategies. When cast in an educational setting, such as chemistry, there might be disadvantages to encouraging all students to remember the shapes of new molecules via mnemonics associating new shapes with unrelated objects. Many students may learn the new molecules better by analysing subset parts of those molecules, for instance, through a key features strategy.
In the current study, the key feature strategy was the most effective strategy, but reporting more than one strategy was also correlated with performance on the decoding task. This is consistent with the notion that spatially skilled people flexibly choose strategies, either cognitively demanding holistic visualization strategies or less cognitively demanding analytic strategies, depending on task demands (Hegarty, 2009).

The mean accuracy rate on the phase 3 visuospatial WM task was 61 per cent. This deserves comment. Since the phase 3 task had three choices, this is well-above random guessing (33 per cent). Although, it may be common practice in cognitive psychology experiments to use tasks above a 90 per cent accuracy threshold, the aim of this study was to use a more typical computer-based visuospatial WM task, similar to those used in educational settings. Sixty per cent accuracy is much more authentic in terms of an instructional situation in which students are learning unfamiliar shapes.

It is intriguing that computer game frequency was significantly correlated with accuracy on the phase 3 visuospatial WM task. It is well-known that computer games have a highly visuospatial component. In some studies, interventions involving computer game play have resulted in increased scores on tests of mental rotation and spatial visualizations (Dorval and Pepin, 1986; Okagaki & Frensch, 1994; Greenfield, 1994). However, in other similar studies, computer game play has not improved mental rotation or spatial visualization (Gagnon, 1985). In the current study, the correlation between computer game play frequency and accuracy on the phase 3 task (visuospatial WM) could have two interpretations: (a) computer game play increases visuospatial WM skills or (b) people who have good visuospatial WM skills are attracted to computer games. This suggests that there are some cognitive similarities between what was done in phase 3 and computer games.

CONCLUSION

Much of the prior research on visuospatial WM has involved relatively indirect sources of data (e.g. factor analysis, brain activation, componential analysis involving latency data, etc.). Such research methods are effective for validating the construct of visuospatial WM and delineating its relationships to other constructs. However, they are less effective for uncovering which strategies are used for visuospatial WM tasks.

The current study used participants’ verbal reports to provide explicit evidence about the role of strategy selection in visuospatial WM tasks. Further, it suggested that strategy choice can be important in performance of visuospatial WM tasks. In this particular study, the key feature strategy appeared to be the most effective strategy in terms of accuracy in a computer-based visuospatial WM task involving unfamiliar shapes. However, in other visuospatial WM tasks with other types of shapes, other strategies may be more effective.

Many other studies of spatial strategies, not only those involving psychometric tests of spatial skills, but also those using mechanical diagrams (Hegarty, 1992), or even academic settings such as chemistry courses (Stieff, 2004), maintain the stimulus shapes in plain view throughout the task. This brings into question whether the shapes are actually being encoded into STM. The current study employed a task where the stimulus shapes were removed from the field of view after being encoded. Thus, it seems more certain that the strategies reported in the current study were actually being used for encoding shapes (or abstractions of shapes) into STM and then subsequently using encodings in WM. More spatial studies of strategy for visuospatial STM/WM should employ interactive tasks and tasks where shapes to be encoded are subsequently removed from view.
An important step in visuospatial WM research is to use participants’ verbal reports, to investigate which strategies are used, and their relative effectiveness, in a variety of different visuospatial WM tasks with a variety of different shapes to uncover other spatial STM/WM strategies that might have educational potential. The target visuospatial WM tasks should include standardized tests of visuospatial WM, interactive computer-based visuospatial WM tasks set in the laboratory, as well as visuospatial WM in authentic educational settings such as college chemistry classes (e.g. Stieff, 2004), as well as recreational and work settings. Such a program of research would shed light on the cognitive processes used in visuospatial WM and highlight how different strategies involve different cognitive processes. Further, the investigation of the relative effectiveness of visuospatial WM strategies would provide valuable information for educators in highly visuospatial disciplines.

REFERENCES


Strategies in visuospatial working memory for learning virtual shapes


