Tetris and Mental Rotation

By
Blaize Michael Kaye

February 2013

Submitted in partial fulfilment of the academic requirements for the degree of Master of Arts in Cognitive Science in the School of Religion, Philosophy, and Classics, University of KwaZulu-Natal.

Declaration – Plagiarism

I, ..., declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   - (a) Their words have been re-written but the general information attributed to them has been referenced
   - (b) Where their exact words have been used, then their writing has been placed inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed

....

Abstract

Research has shown a possible causative link between playing the popular video-game Tetris and improvements in Mental Rotation performance. The aim of the present study was to address a question about an aspect of Tetris expertise that had not yet been factored into any of the existing work on Tetris and Mental Rotation. David Kirsh and Paul Maglio (1994) have shown that skilled Tetris players appear to use *physical* actions as substitutes for, or compliments to, mental operations. This is hypothesised to include physically rotating game pieces instead of Mentally Rotating them. The specific question we sought to address in the present study was whether these physical substitutes for mental operations, which Kirsh and Maglio call *epistemic actions,* have an effect on Tetris' efficacy as a Mental Rotation training task.

In order to address this research question, three groups of subjects were administered tests of Mental Rotation ability before and after a five week training period. The training period consisted of a total of five, hour long, laboratory sessions – evenly spaced across the training period – in which each of the three groups were required to play an assigned video-game. The results showed that a group of subjects (N=13) who received Tetris training on the version of the game that made epistemic actions involving rotation impossible showed no greater Mental Rotation performance gains when their results were compared to a group of subjects (N=13) trained using a Standard version of Tetris. This suggests that the occurence of epistemic actions *does not* have an impact on Tetris' efficacy as a Mental Rotation training task. Further, neither of these two groups showed greater Mental Rotation performance gains than the non-Tetris control group (N=14), a result which suggests that, at least under some circumstances, Tetris training fails to impart Mental Rotation performance gains any greater than what can be expected due to retest effects.

Acknowledgements

First and foremost, I would like to thank my supervisor Professor David Spurrett. Truly, without his guidance, constant encouragement, and shining example I would never have had the confidence to begin, let alone complete, the present work.

I would like to thank my family – Michael Kaye, Lindsay Cook, Helene Kaye, Melody Dean, Tiesa Wood, and Les Burden – for their unflagging support.

Sincerest thanks to my best friend and business partner, Bryan Gruneberg, for helping to make my dream of pursuing postgraduate study in Cognitive Science a reality.

Additionally, I would like to thank my friends in philosophy & cognitive science for their advice and insight – Tyronne Valentine, Wesley Bezuidenhout, Luke Von der Heyde, Adriano Palma, Yolanda Hordyk, and Matt Hitchins.

Thanks, too, to all of the participants in the present study who gave so generously of their time.

Finally, I'd like to thank my wife, Carmen, whose love (and patience) would be impossible for me to comprehend if it weren't for the fact that I feel the same way for her.

Contents

Declaration i

Abstract ii

Acknowledgements iii

Contents iv

1 Introduction 1

2 Literature Survey 3

2.1 Mental Rotation . . . 3

2.1.1 What is Mental Rotation? . . . 3

2.1.2 Sex-dependent differences in MR ability . . . 9

2.1.3 Complexity and Solution Strategy . . . 14

2.1.4 MR Training and Practice . . . 18

2.2 Tetris and Mental Rotation . . . 21

2.2.1 Tetris, individual differences, and MR performance . . . 22

2.2.2 Tetris training and Transfer . . . 28

2.3 Epistemic Actions in Tetris . . . 30

2.3.1 Epistemic versus Pragmatic actions . . . 30

2.3.2 An argument for the existence of Epistemic actions in Tetris 31

2.3.3 Epistemic uses of Rotation . . . 35

2.3.4 Epistemic uses of Translation . . . 37

2.3.5 Do Epistemic Actions increase with skill? . . . 38

3 Aims and Rationale 41

4 Methodology 43

4.1 Research Design . . . 43

4.2 Ethical aspects . . . 44

4.3 Sample . . . 44

4.3.1 Recruitment . . . 44

4.3.2 Group and Subject Allocation . . . 46

4.4 Measurement Instrument . . . 46

4.4.1 Technical specifications . . . 46

4.4.2 Issues with “Web Based Testing” . . . 47

4.4.3 Customisation of WebExp2: Creating a Mental Rotation Test . . . 48
4.5 Apparatus and Materials . . . 51
4.5.1 Computer systems . . . 51
4.5.2 Materials used in orientation, pre-tests, and post-tests . . 51
4.5.3 Customised Tetris implementation used in training task . 52
4.5.4 Control Task - Lemmings . . . 55
4.6 Procedure . . . 56
4.6.1 Orientation and Pre-test administration . . . 56
4.6.2 Experimental Groups - Standard and Modified Tetris . . 57
4.6.3 Control Group . . . 57

5 Results **59**
5.1 Sample Characterisation . . . 59
5.2 Research Question 1 - The effect of training on Mental Rotation ability . . . 60
5.2.1 Data preparation and profile . . . 60
5.2.2 Testing parametric assumptions . . . 62
5.2.3 Research Question 1 core results . . . 63
5.3 Research Question 2 - Does over-rotation of Tetris zoids increase with training? . . . 63
5.3.1 Data preparation and testing Parametric assumptions . . 64
5.3.2 Research Question 2 core results . . . 66
5.4 Research Question 3 . . . 67

6 Discussion **69**
6.0.1 Research Question 1 . . . 69
6.0.2 Research Question 2 . . . 71
6.0.3 Research Question 3 . . . 72
6.1 Issues with the present study and improving subsequent stages of research . . . 74
6.2 Conclusion . . . 75

References **76**

Appendix A - Data analysis tables and figures **82**

Appendix B - Informed consent form **86**

Appendix C - Text advertising research **88**

Appendix D - Briefing text **89**

Appendix E - Mental Rotation test instructions **95**

1 Introduction

Research has shown a possible causative link between playing the popular video-game Tetris and improvements in Mental Rotation performance (Okagaki and Frensch, 1994; De Lisi and Wolford, 2002; Boot et al., 2008; Cherney, 2008).

Tetris is a high speed puzzle game the aim of which is to prevent a wall of shapes – called TETRAZOIDS, or simply, ZOIDS – from collecting at the bottom of the rectangular playing area, sometimes referred to as a WELL, to the point where the accumulated zoids form a structure that reaches all the way to the top of the well. During an EPISODE – defined as the time from when a zoid enters the playing area to when it is placed – the player is able to control the currently descending zoid by rotating it by 90° increments around its local origin or translating it left and right. Players also have the option of dropping the zoid if they are happy with its current orientation and position, thereby ending the episode earlier than if they had waited for the piece to descend at its natural pace. Rotating, translating, and dropping zoids is the total extent of a player's control in the game.

If the player is able to manoeuvre the descending zoids so that they form an unbroken horizontal row across the width of the playing area, the row disappears reducing the overall height of the wall accumulating at the bottom of the well, granting the player points, and extending the length of the game. The game becomes more difficult over time as the rate of the zoids' descent is gradually increased, giving the player less time to place their zoids.

It is the emphasis on zoid *rotation* that has made Tetris seem like a natural training tool for investigating the ways in which video-game play might affect Mental Rotation performance. The present study seeks to address a potential limitation in the existing literature, raised by Boot et al. (2008), by investigating whether Tetris players' supposed use of complexity reducing strategies during training are a factor in their overall post-training Mental Rotation performance.

Kirsh and Maglio's work on Tetris expertise (Kirsh and Maglio, 1994; Maglio, 1995; Maglio and Wenger, 2000) has suggested a number of ways in which intermediate and expert Tetris players may reduce the cognitive demands of Tetris by engaging in what they've termed EPISTEMIC ACTIONS (Kirsh and Maglio, 1994). An epistemic action is any physical action that is performed primarily to reduce the time, effort, or memory that would be required if the problem were tackled mentally. Of particular importance to the present study is a class of epistemic actions that make use of zoid rotation (Section 2.3.3). Kirsh and

Maglio's studies show that intermediate to expert Tetris players *over-rotate* their zoids and argued that the frequency and distribution of these extra rotations are best interpreted as evidence of them serving an epistemic function (Kirsh and Maglio, 1994; Maglio and Kirsh, 1996). The notion that Mental Rotation is required when playing Tetris is an assumption that underlies much of the existing research on Tetris and Mental Rotation performance.[1] Kirsh and Maglio's work suggests that Tetris players can, and do, forgo purely Mental Rotation in favour of rotating zoids physically. If this is the case then there is a possibility that previous studies of Mental Rotation performance and Tetris have, to some extent, misconstrued the nature of the cognitive task posed by Tetris by overemphasising the role of purely Mental Rotation.

In the present research we sought to address this potential oversight in the Mental Rotation and Tetris literature by investigating whether the existence of epistemic actions in Tetris affects the extent to which playing the game improves Mental Rotation performance. Specifically, we wanted to know whether there would be a measurable difference in post-test Mental Rotation performance between a group of participants trained using a standard version of Tetris and a group of participants who were trained using a version of the game modified in such a way that complexity reducing strategies are made difficult or impossible.

To this end, we conducted an experiment that began with the administration of a test of Mental Rotation ability to three groups of subjects. After completing this pre-test, each of the groups were assigned video-games that they were required to play for at least 5 hours over a 5 week period. Two of the groups were assigned different versions of Tetris, one of which had been modified to prevent subjects from over-rotating their zoids while the other was a standard version of the game that allowed over-rotations.[2] The third group was assigned a control task, a video-game that did not involve rotation. At the end of their training period we administered a second test of Mental Rotation ability.

[1]For explicit statements that Mental Rotation is a requirement for Tetris play one can look at, for example, Okagaki and Frensch (1994),De Lisi and Wolford (2002), and Sims and Mayer (2002).

[2]In this context we can define "over-rotation" of a zoid to be any number of rotations greater than, or equal to, the number of rotations required to get the zoid back into the orientation in which it entered the playing area.

2 Literature Survey

2.1 Mental Rotation

2.1.1 What is Mental Rotation?

When presented with the pair of items labelled A in figure 1 most people, assuming that they are not distracted and are given enough time, will be able to tell that the object shown on the left is identical to the one on the right despite the fact that the object is shown from two different perspectives. Furthermore, the time a person takes to recognise that the two images represent the same object seems to be related to the magnitude of the angular displacement of the object as it's presented from the two different points of view (Shepard and Metzler, 1971). It is the ability to perform this, and similar, kind of operations unaided by the use of external manipulations or props that is referred to as Mental Rotation.[3]

It is important to begin with a clear behavioural characterization of MR because, despite almost forty years of research, we do not yet completely understand the nature of the processes and representations involved (Pylyshyn, 2002). There is, for example, strong evidence to suggest that there may be more than one process underlying the behaviour that we're trying to explain (Geiser et al., 2006). However, there is a fairly broad consensus about certain aspects of MR that emerged out of the pioneering work of the 1970s and early 1980s. In the present section we will be reviewing the work that established MR as a psychological phenomenon and settled, to some extent, some of its important properties (Linn and Petersen, 1985).

The scientific study of Mental Rotation begins with the seminal work of Shepard and Metzler. Their paper "Mental Rotation of Three-Dimensional Objects" (Shepard and Metzler, 1971) is considered a landmark in the history of cognitive psychology. Not only did it introduce an important psychological phenomenon, Mental Rotation, into the scientific study of the mind, but their paper also showed that with some ingenuity it was possible to investigate, quantitatively, an entire class of phenomena that had primarily been subject to subjective, introspective investigation, namely mental imagery (Goldstein, 2008).

Their experiment involved presenting their subjects ($N = 8$) with pairs of two-dimensional projections of three-dimensional objects, examples of which are

[3] For the sake of readability, we will often abbreviate "Mental Rotation" to MR.

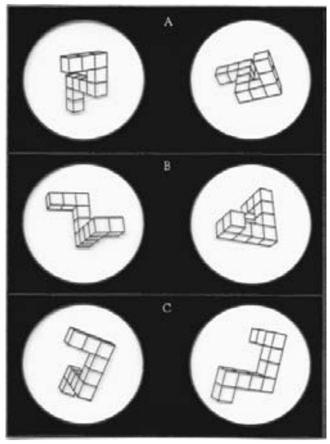

Figure 1: Examples of the shapes used in Shepard and Metzler's 1971 experiment (Shepard and Metzler, 1971)

shown in figure 1. Each trial presented a pair of images in which both the target and comparison objects were identical, or in which the comparison object was an isomer (mirror image) of the target object.[4] In the latter case the mirror image of the target was used as a comparison object so that the participants weren't able to rely on any local features distinctive to either object to distinguish them from one another (Shepard, 1978). The target and comparison objects were always shown at different orientations with the comparison object rotated around its local origin at multiples of 20°, either in the image plane or in depth, that is, around its z or y-axes respectively. On presentation of an experimental trial a timer would start and subjects were required to indicate, as quickly and as accurately as possible, whether the images represented identical or different objects. A total of 1600 trials were presented to each subject and these were split evenly between "same" and "different" trial pairs.

[4]Terminological note: Although not applicable to all MR experiments, we will generally refer to the image or object that appears in an upright or standard orientation as the TARGET item, and the image or object that is to be "rotated" or that appears at a non-standard orientation as the COMPARISON item.

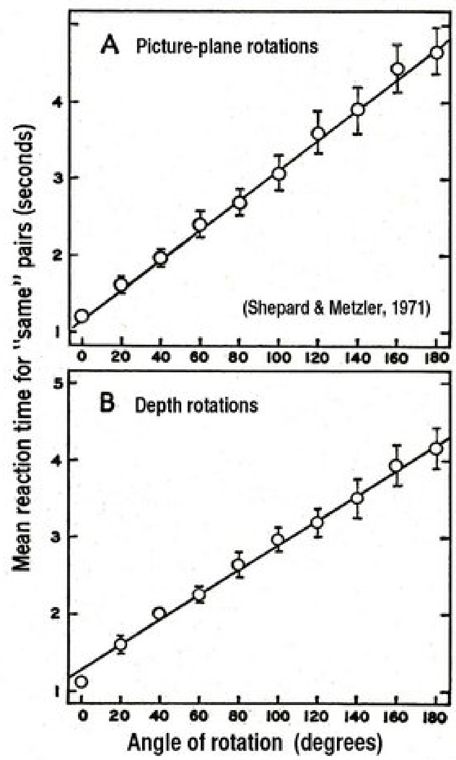

Figure 2: Graph of Response Times by angular disparity (Shepard and Metzler, 1971)

Figure 2 shows the results of plotting response times against the angle of rotation for "same" trials in which subjects answered correctly for both rotation in the picture plane (the upper graph) and in depth (the lower graph). What was, and still is, striking about these graphs is their linearity. Both in-depth and in-plane graphs show a clear linear relationship between angular displacement and Response Time.[5] The researchers calculated polynomial regression lines for all subjects' data individually and in every case found a highly significant linear relationship ($p < .001$) but no significant quadratic or higher order effects (Shepard and Metzler, 1971). When asked to provide an introspective account of how they performed the task, all subjects reported that they had imagined a representation of the comparison object and then "mentally rotated" this object to bring it into congruence with the target. Shepard and Metzler (1971) note

[5]Response Time, or Reaction Time, is often abbreviated as RT

that introspective reports are not reliable data, but the subjects' interpretation is in line with their hypothesis that the process underlying MR involved a kind of "functional equivalence" (Sternberg and Sternberg, 2011) with rotating objects in physical space. Shepard (1978) fleshes out this functional equivalence by arguing that, while there need not be so-called *first-order isomorphic* relationships between the neural states underlying a process like MR – that is, for example, there need not be a "table like" state somewhere in the brain in order for the brain to represent a table – there has to be some more *abstract second-order isomorphisms* "in which the functional relations among objects as imagined must to some degree mirror the functional relations among those same objects as actually perceived" (Shepard, 1978). The second-order isomorphic relationship at play in instances of Mental Rotation is then, at the very least, that whatever represents the comparison image in mind is subject to a continuous transformation around its origin, through all intermediate orientations, until it is brought into correspondence with the target image.

A number of researchers questioned the assertion that the subjects were rotating the images holistically (Just and Carpenter, 1976; Hochberg and Gellman, 1977). Just and Carpenter (1976), for example, tried to show that the linear relationship between angular displacement and Response Time can be explained without the need to posit a holistic transformation of an image-like mental representation by arguing, from eye tracking data, that MR is a piecemeal process whereby segments of the comparison item are converted, or encoded, into amodal, symbolic representations which are serially transformed through a number of discrete steps. Just and Carpenter observed, in an experimental setup that essentially replicated Shepard and Metzler (1971), the number of times a subjects' eye-fixation switched between images was, like Response Time, linearly related to the overall angular displacement between target and comparison images. From these data they argued that because eye-fixations could correspond to the discrete transformation of a mathematical representation of a section of the comparison image, the overall increase in Response Time need not be explained by the greater angular distance that needs to be swept through as a mental-image is rotated, but could rather be explained by the fact that there is simply a linear growth in the number of computations required.

As Just and Carpenter themselves point out, their results are difficult to generalise to all instances of Mental Rotation. There are a number of tests of MR ability that do not rely on the side-by-side presentation of target and comparison images. For example, Cooper and Shepard's (1973) experiments

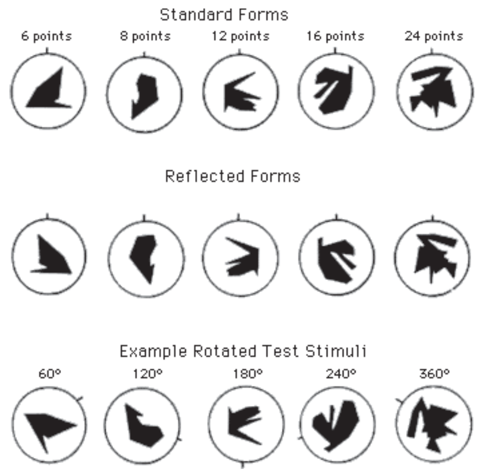

Figure 3: Examples of polygons used in Cooper (1975)

presented subjects with familiar stimuli, a number or letter of the alphabet, or its mirror image, that had been rotated around its local origin. In this case the target image didn't need to be shown because the subjects knew what the images looked like at their standard orientations. Subjects were, once again, required to indicate as quickly and accurately as possible whether the comparison image was merely a rotated version of the target, or a rotated *and* reflected version of the target. Response times for the experiment showed that the same linear relationship with angular displacement as the side-by-side protocol even though there was no target image for comparison.

Experiment 2 of Cooper's 1975 paper presents further evidence for the notion that a "mental-image" is rotated holistically. The experimental task resembles earlier MR experiments in that subjects are required to distinguish images, in this case more or less complex polygonal shapes, from their mirror images. The difference is that in Cooper (1975) subjects are asked to mentally rotate the target image into a particular orientation *before* the comparison image is presented. Subjects ($N = 8$) were first trained to recognize certain of the polygonal shapes (see figure 3) at fixed orientations, these shapes and orientations then represented the "standard" or canonical versions of the images. The experimental phase consisted of trials made up of three parts. First, one of the canonical items, at the learned orientation, would be shown to the subjects for 3 seconds. Second, the subjects would be presented with information about the orientation in which the comparison image would appear. This was achieved through displaying an arrow pointing in the direction that the top of the target image would have to be rotated to in order to bring it into congruence with whatever image appeared as comparison. Subjects were told to ready themselves for the discrimination task and indicate, through the press of a button, when they felt prepared, at which point the comparison image would appear. This was the third stage where subjects were required to indicate, as quickly and accurately as possible, whether the comparison image was identical to the target image by saying "S", for "similar", or "R", for reflected, into a microphone that would stop the timer that had been running since the presentation of the comparison images. Two response times were therefore recorded. First there was the time interval between the presentation of the orientation cues and the pressing of the button to indicate readiness, call this *preparation time*. Second, there was the time interval between the presentation of the comparison image and the subject indicating whether the two images were similar or reflected, call this *discrimination time*. The results of the experiment were unambiguous – *preparation time* showed a clear linear relationship with the difference in orientation between target image and comparison image, while *discrimination time* was more or less constant regardless of the angular disparity between the two. Cooper argued that these results show that subjects are creating and maintaining a mental-image of the target image that is then mentally rotated to the indicated orientation. The speed of this rotation – being an analogue of physically rotating an object – should have some upper limit[6], which would account for the characteristic linear relationship between angular disparity and the subjects signalling that they are prepared. The *discrimination time* is, then, taken to be constant because the mental image of the target has *already* been rotated to the appropriate orientation in working memory, meaning that the only thing required of the subject during *discrimination time* is to tell whether target and comparison images are identical.

While both Cooper and Shepard (1973) and Cooper (1975) demonstrate that the rotational transformation of mental-representations is able to take place without the need for the target and comparison images being visually present, Cooper (1976) goes much further in establishing that the mental-representation of the comparison image *actually* passes through a series of intermediate orientations while being brought into correspondence with the target. This later

[6] Shepard and Metzler (1971) calculate the speed of Mental Rotation to be around 60°per second. Subsequent studies have found that this varies substantially across type of stimulus and amount of practice.

study used 6 of the same subjects, as well as the stimuli, from (Cooper, 1975). Using data from the previous study allowed the researcher to calculate rates of Mental Rotation for every subject. Each trial then consisted of presenting the target image, one of the previously learned stimuli at the learned orientation, for two seconds. The screen was cleared and the subject was then presented with a blank circular field during which time they were required to mentally rotate the target in a clockwise direction. After a preset interval the comparison image, again identical or the mirror image of the target, would appear at one of 12 equally spaced orientations in the circular field. Given that subjects' rate of Mental Rotation was estimated upfront, Cooper hypothesised that it should be possible to predict at which orientation the mental image would be when the comparison image appears. In "probe-expected" trials, the comparison image would appear at exactly the orientation that was predicted by the subjects' rate of MR, while in "probe-unexpected" trials the comparison would appear at some other orientation. The results of the experiment supported Cooper's hypothesis. The reaction time for "probe-expected" trials were almost constant regardless of the orientation at which the comparison image appeared while reaction times for "probe-unexpected" trials increased linearly with the angular disparity between the comparison image's actual orientation and the predicted orientation of the mental-image given the subject's speed of Mental Rotation.

(Cooper, 1976) manages to capture the two important aspects of the early work on MR. Firstly, it demonstrates the clear relationship between MR response times and angular disparity. Secondly, it demonstrates, quite persuasively, that under at least some circumstances the representation that underlies the MR phenomenon passes through successive orientations until it is brought into correspondence with the comparison image.

After the early work of Shepard and his students had established the existence of MR – to the satisfaction of a large part of the psychological community at least (Linn and Petersen, 1985) – the investigation of MR has largely turned towards those factors impacting individual performance in tests of Mental Rotation ability.

2.1.2 Sex-dependent differences in MR ability

In a meta-analysis of early work on sex-dependent differences in spatial ability, Linn and Petersen (1985) found that, while males sometimes outperform females on a number of measures of spatial ability, the largest and most persistent difference in visual-spatial ability is found in tests of MR performance. Subsequent studies have tended to corroborate these findings (see, for example (Peters et al., 1995; Masters, 1998; Peters and Battista, 2008)). It is because of this persistent and often substantial sex-dependent difference in performance that a relatively large part of the work on MR has been devoted to investigating the conditions under which this difference manifests. That is, rather than merely investigating absolute differences in MR performance for groups of subjects from different Socio-economic groups (Levine et al., 2005), cultures (Mann et al., 1990), or academic programmes (Peters et al., 1995), it is often the way that the differential performance of males and females varies between those groups that is a central point of interest. From a theoretical perspective, an understanding of these sex-dependent differences affords us deeper insights into the cognitive differences between males and females, if indeed there are any. From a practical perspective, there is interest in developing effective interventions to reduce the apparent sex-dependent difference in MR performance. With regards to this latter point, it has been suggested that visual-spatial skill may serve as a mediating factor in gender-based mathematics differences (Casey and Nuttall, 2001) and that effective methods of improving visual-spatial performance of females might have an impact on female involvement in subjects and careers with a mathematical foundation (Cherney, 2008), in which females are still largely under-represented (Ceci et al., 2009).

There is evidence to suggest that a sex-dependent difference in spatial performance in general and MR in particular, emerges remarkably early in childhood development. For instance, Moore and Johnson (2008) used a habituation method to study whether there was a sex difference in MR performance in 5 month old infants (N=40, 20 female). They showed their subjects animations of 3-D Shepard-Metzler style blocks rotating backwards and forwards through 240°until the infants were habituated. They then showed either the habituation object rotating through the previously unseen 120° or they presented the subjects with a novel object, namely, the habituation object's mirror image. The results were that the male subjects looked at the novel stimulus significantly longer than the habituation object at an unfamiliar angle ($p < .001$) while the female subjects looked at both habituation and novel objects approximately equally. The authors reasoned that the males' preference for the novel stimulus was evidence that they were recognising the habituation objects at unfamiliar angles supporting the notion that the male subjects, but not the female subjects, were engaging in some kind of MR process. Similar sex-dependent MR performance differences have been found in 3 month olds (Moore and Johnson, 2011) and with 2-D stimuli in 2 to 4 month old infants (Quinn and Liben, 2008). While these studies do provide evidence for a sex-dependent difference in spatial abilities, because they focus solely on Mental Rotation, they do not necessarily demonstrate that male infants perform better at MR *in particular* as compared to other visual-spatial abilities. The performance difference noted in these particular studies may be a manifestation of a more general male visual-spatial performance advantage in infants, and indeed, there is reason to believe that the larger MR *specific* performance difference emerges only much later in development. Levine et al. (1999), for instance, found that *male* pre-schoolers between the ages of 4 to 6 performed significantly better on two different tests of visual-spatial ability. The first of the tests required subjects to mentally transform, including mentally rotate, shapes while the other was the mazes sub-test of the 1989 revision of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-R) which requires the subject to solve a series of paper and pencil mazes of increasing difficulty. The latter test was chosen *primarily* because it seems not to require MR. As was mentioned, males performed significantly better on *both* tests than females but there was no evidence to suggest the particularly robust difference in MR that is characteristic of studies of adult subjects (Linn and Petersen, 1985). This cannot, however, be taken as conclusively demonstrating that such a robust different does *not* exist in preschoolers though. This is because, as Levine et al. (1999) point out, it is extremely difficult to design a test of MR ability that is equally appropriate for children, adolescents, and adults. This is a particularly important point because – as we discuss below – even in adults the magnitude of the sex-different in MR performance seems to vary with the *kind* of test being administered (Peters et al., 1995; Jansen-Osmann and Heil, 2007).

There are several considerations that make the investigation of a possible biological account of these sex-dependent differences in MR performance a natural move. Firstly, as we have just discussed, the difference seems to emerge quite early in development. Secondly, as mentioned above, the difference has been shown to hold cross culturally. Geary and Desoto (2001), for instance, investigated whether similar patterns of sex-dependent differences would hold between adolescents in China and the United States, seeking to supplement and extend the results of an earlier study (Mann et al., 1990) that had demonstrated the, now familiar, male performance advantage in Mental Rotation in Japanese adolescents. In their Study 1, their subjects, a group of Chinese (N = 20, 20 female) and American (N = 66, 42 female) undergraduate students, were administered a series of spatial tests including the Card Rotations Test (CRT) and the Cube Comparisons Test (CCT), both of which are from the Educational Test Service battery of factor-referenced tests, as well as the Vandenberg-Kuse Mental Rotation Test (MRT) (Geary and Desoto, 2001). The Card Rotations Test is a pencil and paper task in which subjects are presented with a series of paired simple polygons and asked to indicate whether the shapes are the same, and at different orientations, or different shapes entirely. The Cube Comparisons Test (CCT) is a test of 3-D MR ability that presents subjects with 2-D projections of two cubes. The cubes are drawn in such a way that three of their faces are visible, with each face showing a letter of the alphabet. Subjects are then required to determine whether either of the cubes could be rotated so that its faces match the faces of the other. The Vandeberg-Kuse MRT uses the Shepard-Metzler blocks in a pencil and paper MR task that requires the subject to Mentally Rotate in the picture plane, in-depth, or both at once making it the most difficult of the three. Subjects taking the MRT are presented with a target object, a Shepard-Metzler 3-Dimensional cube form, and four comparison objects and are asked to tell, as quickly as possible, which of the comparison objects are identical to the target. The MRT is typically time-limited and dependent variables of interest are most often the number of correct responses, the number of errors, or some combination of the two. The Vandenberg-Kuse MRT is a particularly important instrument in studies of sex-dependent differences in MR because across all tests of Mental Rotation ability it provides the largest, and most consistent, performance difference (see (Peters and Battista, 2008)). In (Geary and Desoto, 2001) participants were given 3 minutes, for each of the three tasks, to answer as many items as they could – final scores for all three tasks were the number of items the subject scored correctly minus the number of incorrect items. The researchers ran a 2 (nation) by 2 (sex) ANOVA on the scores for each of the three tests. They found no significant main effects or interaction effects for the Card Rotations Test, but found a significant ($F(1{,}102) = 15.83$, $p < .001$) main effect for nation and interaction effect for nation by sex ($F(1{,}102) = 6.62$, $p < .05$) for the Cube Comparisons Test (Geary and Desoto, 2001). The analysis of the MRT results showed a main effect for sex ($F(1{,}102) = 14.55$, $p < .001$) favouring males in both nation groups, and showed no significant main effect for nation ($F(1{,}102) = 2.15$, $p > .10$) or nation by sex interaction ($F(1{,}102) = 1.55$, $p > .20$).

Geary and Desoto (2001) favour a biological explanation of this performance difference and in support of their account they point out not only that the difference has been found in heterogeneous cultures across the US and Europe (Voyer et al., 1995), in Africa (Amponsah and Krekling, 1997), and now *again* in East-Asia, but further, appeal to the fact that MR performance has been shown to be related to levels of sex hormones (Hooven, Chabris, Ellison and Kosslyn, 2004).

However, as Geary and Dosoto (2001) themselves point out, an account of the sex-dependent performance difference that *simply* appealed to biology would fail to account for all the facts, environment is clearly an extremely important factor in the development of visual-spatial ability. An interesting example is the 2005 paper in which Levine et al. present the results of a longitudinal study that ran over the course of two years with the intention of investigating what, if any, effect Socio-economic Status (SES) has on this gender gap in spatial ability. SES status was assigned at a school level on the basis of census-track data for Illinois. A total of 547 students were recruited for the experiment, with male and female participants being approximately equally represented across three SES groups – high, medium, and low. Testing consisted of administering an aerial-map task in which participants were asked to draw correspondences between aerial photographs of an area and a map of the same area, a mental rotation task based on the Spatial Relations subtest from the Primary Mental Abilities (PMA) Readiness Level, and a syntax comprehension test (Levine et al., 2005: 842). Given the persistence of sex-dependent differences across an extensive number of studies of MR performance, Levine et al. expected to see performance differences manifest in the spatial tasks but not the language task. This expectation was mostly borne out by their results with an exception, namely, that the expected differences in spatial skill held only for middle and high SES subjects (Levine et al., 2005). Low SES male and female subjects, however, failed to show any significant differences in their performance on the aerial-map and mental rotation tasks. That is, in Levine et al's study, the gender gap is virtually non-existent for the low income group.

The researchers posit two possible explanations for their findings. The first starts with the observation that, generally, the sex-dependent performance difference manifests itself in the more difficult test items, as was demonstrated in Geary and Desoto (2001) above – if both male and female low SES group subjects had, for some reason, failed to succeed in answering the more difficult questions then any sex-dependent difference wouldn't manifest in the data even if a difference did in fact exist. However, further analysis of their data seems not to support this hypothesis, for example, a difference in spatial ability between males and females in the low SES group failed to manifest in the subset of data where performance across all three groups was comparable for spatial tasks, while the difference persisted for the higher groups. A second possible explanation for the results, and the one the researchers (and their data) seem to favour, is the notion that it is "differentially high level[s] of engagement in the kinds of activities that promote the development of spatial skill[s]" (Levine et al., 2005: 884) that causes the gender gap in spatial ability, and that these kinds of activities (playing with particular toys, freely exploring their neighbourhoods, etc.) might not be as readily available to males from low SES groups as they are to males in other SES groups or are equally available to both males and females in low SES groups.

2.1.3 Complexity and Solution Strategy

There is evidence to suggest that *solution strategy* is an important determining factor in subjects' performance in Mental Rotation tasks. The possibility of distinguishing between MR solution strategies was explicitly raised in Cooper and Podgorny (1976) in the course of their investigating a theory that if the representations underlying MR are image-like then they would likely be rotated holistically and if they are language-like/propositional then they would likely be rotated using a piecemeal process. They claimed that while both possible types of representation (image-like or language-like) could explain Response Times increasing with increasing angular discrepancies, it should nonetheless be possible to distinguish between them. In particular they argued that if MR is accomplished in a piecemeal fashion then MR Response Times should become slower as the image to be rotated becomes more complex. That is to say that if one performs a linear regression on the data representing subjects' Response Times against angular discrepancy, the *slope* of the regression line – which can be interpreted as representing the Mental Rotation process itself – should vary with the complexity of the figures being rotated. This increase in Response Time with complexity would be attributed to the fact that with a piecemeal MR process transforming complex figures simply requires a greater number of discrete operations than translating simple figures.[7] If, on the other hand, MR is accomplished using a holistic process then we should expect that the slope of

[7]This argument fails to take into account the possibility that multiple discrete operations could take place in parallel (Smith and Dror, 2001).

the regression line should remain constant regardless of the complexity of the figures being rotated.

Cooper and Podgorny found no evidence showing that their subjects' Response Times varied with stimuli complexity and from this argued that their data therefore support the notion that MR processes are holistic and that the representations underlying the process are at the very least not of a "simple class of propositional models" (Cooper and Podgorny, 1976: 505). These results and inferences are slightly problematic because, firstly, their study used the same subjects that had participated in Cooper (1975) and, secondly, their task required their subjects to Mentally Rotate the same stimuli used in the earlier study. At best Cooper and Porgorny's (1976) results suggested that subjects who have been trained in MR, tasked with rotating highly familiar stimuli, demonstrate constant MR performance regardless of stimulus complexity.

That MR performance does *in fact* vary with stimulus complexity has been demonstrated a number of times since Cooper and Podgorny's (1976) study (see, for example, Pylyshyn, 1979; Folk and Luce, 1987). Bethell-Fox and Shepard's (1988) study is particularly noteworthy in that it addresses the problems with Cooper and Podgorny's paper directly. Specifically, Bethell-Fox and Shepard were interested in readdressing the question of whether complexity of stimuli had any significant impact on MR performance, and if any differential performance due to stimulus complexity was discovered, whether extensive practice with the stimuli would reduce or eliminate this difference. Stimuli for the two experiments presented in their study were made up of 3x3 grids in which a number of the cells were filled in forming completely asymmetric patterns, this was to ensure that the patterns were unique through all 8 orientations determined by 90 degree rotations and reflection. Each of the figures were assigned an inverse measure of complexity[8] that Bethell-Fox and Shepard term *figural compactness*[9].

Their Experiment 1 required 8 undergraduates to perform a MR task similar to that of (Cooper, 1975). Each trial in the procedure had three phases. Firstly, when a subject was ready for the trial to begin they would press a button and

[8]In fact, two *different* measures of complexity were assigned to each of the stimuli. In addition to *compactness* Bethell-Fox and Shepard (1988) assigned each figure a measure of complexity based on the number of unattached groups of filled in blocks in the 3x3 matrix. This latter measure was found to explain inspection, rotation, and comparison time as well as, but not better than, *figural compactness*. This means that, statistically speaking, the choice between one measure over the other with regards to Bethell-Fox and Shepard's study's stimuli is arbitrary.

[9]Figural Compactness is defined as $Compactness = \frac{\sqrt{Area}}{Perimeter}$ where $Area$ is given by the number of filled in blocks, and $Perimeter$ is given by the number of exposed block sides (Bethell-Fox and Shepard, 1988).

the target image, one of the 3x3 matrices, would be presented at one of 8 possible orientations. Second, the subject would indicate, by pressing a button, that they had sufficiently studied the matrix which would then be replaced with a rotational cue, indicating that the subject should Mentally Rotate the matrix either 90 degrees or 180 degrees clockwise or anti-clockwise. Finally, once the subject had performed the Mental Rotation they were shown a comparison image and the subject would be required to indicate, by pressing one of two buttons, whether it was the same or different as the target. This, like (Cooper, 1975), gave three distinct measures. The first phase was interpreted as representing the time it took for the subjects to encode the stimuli, the second phase represented the time taken to actually rotate the matrices, while the third phase provided a measure of the time taken to compare target and comparison images. In addition to their subjects' response data displaying the standard "Mental Rotation Effect" (increasing RT with angular displacement), and contrary to (Cooper and Podgorny, 1976) , Bethell-Fox and Shepard found clear evidence that complexity has a statistically significant effect on MR performance. Specifically, they found that the differences in Response Time attributable to stimulus pattern was statistically significant for the inspection phase ($F(17,\ 102) = 7.10$, $p < .001$) rotation phase ($F(17{,}102) = 7.99$, $p < .001$) and the comparison phase ($F(17{,}102) = 7.99$, $p < .001$). As the complexity of the figures (given by their *figural compactness*) increased, there was a corresponding increase in RT across all three phases. On the other hand, in a result that explains the earlier findings by Cooper and Podgorny (1976), Bethell-Fox and Shepard's Experiment 2 revealed that the effect of stimuli complexity disappeared for stimuli with which subjects were given extensive MR training, although they still displayed a linear relationship between RT and angular disparity.

It is clear that any suggestion that Mental Rotation has to be an *exclusively* holistic process is going to be a non-starter because there *has to be* some upper bound on the complexity of figures that are able to be rotated holistically. For instance, it may be that images of human faces are too complex for us to Mentally Rotate all at once. Takano and Matia (2006) suggest that a failure of holistic Mental Rotation, due to the complexity of human faces, may be illustrated by, and could help explain, Thompson's "Thatcher Illusion" (Thompson, 1980). This is the phenomenon where an inverted image of a human face, the eyes and mouth of which have been manipulated into their standard, upright, orientation, looks almost normal but when the image is viewed the right way up (with the mouth and eyes now inverted) the changes that have been made to the face become much more apparent.[10]

It seems that often a holistic rotational strategy may be more efficient than a piecemeal strategy. This includes instances where the images to be Mentally Rotated contain what Hochberg and Gellman (1977) call *landmark features*. Landmarks are salient points or features that serve as orientating markers or points on stimuli too complex to be apprehended in a single glance. Hochberg and Gellman (1977) performed a standard Shepard-Metzler side-by-side MR experiment with stimuli designed with or without landmarks. MR performance was shown to be substantially better with landmark rich stimuli. More precisely, for a linear regression of subjects' Response Times on Angle of Rotation both the intercepts – interpreted as time to encode stimuli – were significantly shorter ($p < .001$) and slopes – interpreted as rate of Mental Rotation – were significantly smaller ($p < .001$) for the landmark rich stimuli than for landmark poor stimuli.

Presently the evidence suggests that humans are able to make use of both piecemeal and holistic/analogue MR strategies (Pinker, 1998), and that which strategy is used in any instance of MR is determined by a number of factors, not only stimuli complexity. For instance, a recent study by Dror et al. (2005) found evidence suggesting that older and younger adults tend to make use of different MR strategies. They administered a Mental Rotation task with two classes of stimuli, complex and simple, to a group of 16 younger adults (mean age = 18.1 years, SD = 1.4 years) and 16 older adults (mean age = 69.9 years, SD = 7.8 years). Both age groups demonstrated a Mental Rotation effect, but only the younger adults' rates of Mental Rotation changed in response to increased stimuli complexity, suggesting that the younger adults were prone to using piecemeal rotation. Dror et al. (2005) argue that because a piecemeal strategy requires multiple mental transformations of complex representations it is the more mentally taxing of the two. A holistic strategy, although not as flexible as a piecemeal strategy – as mentioned above, there seems to be an upper limit to what can be rotated holistically – is a simpler process and presumably less mentally taxing because it only requires a single, continuous transformation of a single, unified representation. Further, they argue that older adults seem to be adopting the simpler, but less flexible, of the two strategies as a dynamic compensation for generally declining cognitive resources and capacities.

[10] MR stimulus complexity is only one possible factor in, or explanation of, the "Thatcher Illusion". For further information about how inversion of distorted facial image impairs processing, see Bartlett and Searcy (1993).

2.1.4 MR Training and Practice

Mental Rotation performance is exceptionally responsive to practice. Simply retaking a test of MR ability is often enough to elicit a statistically significant improvement in Response Time (Peters et al., 1995). Some of the most useful data comes courtesy of Robert Kail's research into the effects of practice on MR ability (Kail, 1986; Kail and Park, 1990). Kail & Park (1990) set out to verify and extend the results of the earlier study (Kail, 1986) in an investigation of the effect that massive amounts of practice on an MR task – over 3000 practice trials – would have on the MR ability of children and adults.

The study had two aims. Firstly, they were interested in determining the shape of the function that best described the relationship between amount of practice and MR performance in order to help determine whether the process underlying MR improvement is the same in adults and children. Secondly, they were interested in investigating the breadth of transfer from the training task to a second, different MR task, as well as to another speeded task that did not require Mental Rotation. An experimental and control group, each consisting of 8 adults and 8 children, were administered these three tasks as pre- and post-tests. In the first MR task subjects were required to indicate, quickly and accurately, whether letters (F, G, P, or R) presented at different orientations were mirror-images of their standard presentation or not. This first task was used to gauge the effects of practice on MR ability – between pre- and post-tests the experimental group were exposed to an additional 3000+ practice trials of this task, all of which were recorded. The second MR task, used to measure near transfer of training, required subjects to say whether differently oriented letter-like items from the Primary Mental Abilities (PMA) test, presented using the side-by-side paradigm, were identical or mirror-images. The third pre- and post-test task was a memory search in which subjects are presented with an array of between 1 and 5 digits which they were required to study for approximately 4 seconds, after which they were presented with a single digit and asked to specify, as quickly as possible, whether the single digit had appeared in the array.

For the MR tasks, a linear regression was run on every subject's data in order to calculate two parameters related to performance. The first, given by the x-intercept of the best fitting line, represented the average time it took the subjects to encode the comparison image, compare it with the target, and respond. The second parameter, representing MR performance *per se*, is then associated with the slope of the line – this is taken to represent the rate at which the subjects are able to rotate the mental-image. Analysis revealed that although both groups experienced improvement, training significantly improved the experimental group's performance in the MR task involving letters across the two sessions ($p < .05$). Within the experimental group, both adults and children experienced significant decreases in their intercepts, suggesting that the speed of encoding, comparing, and responding had increased significantly with training. Children experienced a significant ($p < .05$) improvement in their MR rates, represented by the slope of the regression line, while adults' rates of rotation remained relatively stable across pre- and post-tests. Furthermore, there was no evidence that the training received in the one MR task had any impact on the rotation of letter-like images from the PMA, or any significant impact in performance in the memory search task. These latter two results suggest that MR training is highly task specific (see also Sims and Mayer, 2002: discussed below).

The results showing improvement after practice are not in themselves surprising – the real value of Kail and Park's study lies in the data that they collected during the training sessions showing the improvement in MR ability as a function of practice. As mentioned above, from these data the researchers were able to perform a curve fitting exercise with the aim of identifying the function best describing this relationship. As with the earlier study (Kail, 1986), they found that the data was better fit by hyperbolic and power functions than exponential functions.

Figure 4 presents a scatter plot of these data, along with the best fitting hyperbolic curve for both the children's and the adults' responses over time. As can be seen, the children begin with a slower MR rate but quickly reach asymptotic levels of performance comparable to those of the adults.[11] Importantly, Kail and Park's analysis revealed that the functions were an adequate fit for both sets of data even when the functions' parameters were shared between the two. This was taken to show that the process underlying improvement in MR performance is identical in children and adults, the only difference being that adults' begin closer to asymptotic performance than children. Kail and Park's results clearly indicate that there is some kind of drastic improvement

[11] For the present study it is useful to consider how much Tetris would need to be played for a Subject to be exposed to an approximately equivalent number of MR "trials". For our purposes, let us consider each Tetris Episode (defined above) to be equivalent to one of Kail's MR trials. The subjects who participated in the present research had an average of 79 episodes per game (excluding Square-shaped zoids, which do not need to be rotated). Very roughly this means that, for our subjects at least, about 38 games of Tetris should provide an equivalent number of practice trials as Kail and Park's experiment (1990).

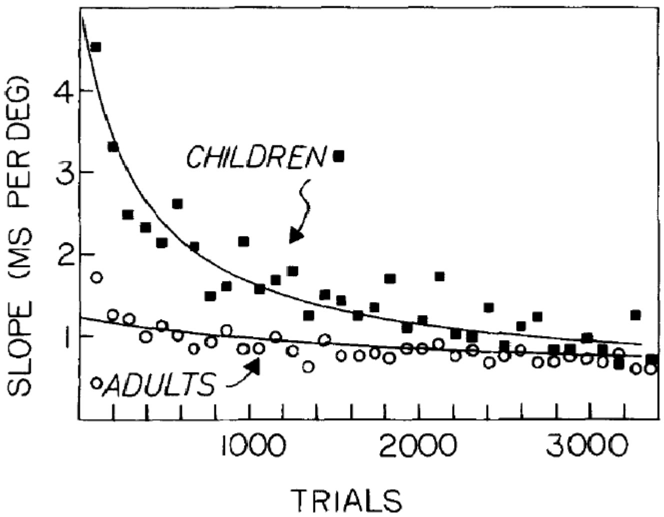

Figure 4: MR performance improvement with practice in children and adults (Kail and Park, 1990)

in response times for the training task, however, they argue that if it were the process of rotation itself that was being improved, then there should have been transfer to the PMA based task. Instead, they suggest that as subjects become more familiar with the stimuli through training, they build up a storage of representations that allow the subject to respond without needing to actually Mentally Rotate the comparison image. Adults, then, may have started out closer to asymptotic levels of performance because they had more stored representations of letters at different orientations (Kail and Park, 1990: 243). Note that they are not suggesting that there is no MR process, only that the process of Mentally Rotating images can be short circuited if the subject has a stored representation of the image. Kail and Park refer to this kind of account as an *instance-based* explanation of MR improvement.[12] This kind of instance-based account is supported by Tarr and Pinker's (1989) study of MR in which they found that, after practice, subjects' Response Times for MR tasks, on learned items only, did not always show the characteristic relationship with angular

[12] *Instance-based* accounts of the change from unskilled to skilled performance are often contrasted to *process-based* theories which, generally, postulate that the improvements in behaviours like MR are due to optimisations in the underlying processes themselves – redundant steps are thought to be eliminated, processes may be used more effectively, discrete processes are bound together into larger units (Heil et al., 1998).

discrepancy, but that RTs were more or less constant for the trained stimuli – again suggesting that the process of Mental Rotation was able to be bypassed by using stored representations.

Although inconclusive, there is some evidence to suggest that not all MR performance improvements are *instance based,* that is, the MR process *itself* may be able to be improved. A clear indication of instance based improvement to MR is that there is very little transfer of RT improvement from trained to novel stimuli. In a very recent study, Wiederbauer and Jansen-Osmann (2008) were able to demonstrate that children displayed statistically significant improvement to their MR performance for both highly familiar *and* novel stimuli after a period of training with a computer based task requiring the use of a joystick to manually rotate 2-D pictures of animals, presented side-by-side, into congruence with one another. These findings are in contrast to an earlier study involving the same researchers (Wiedenbauer et al., 2007) that used a 3-D version of the same experimental design with adult subjects and found no real evidence of broad transfer from training to novel stimuli (see also (Heil et al., 1998)).

We will return to the question of the training of Mental Rotation when we deal with work using Tetris to study Mental Rotation below.

2.2 Tetris and Mental Rotation

Video-game play is a hugely popular pastime. It is estimated that somewhere around 60% of Americans are regular game players (Green and Bavelier, 2006) and, perhaps even more surprisingly, in 2009 the British public spent more money on video-games than they did on visiting the cinema or purchasing DVDs for home viewing combined (Wallop, 2009). More importantly, with the world-wide proliferation of cellphones as low cost computing devices able to run relatively sophisticated computer programs, the potential audience for this form of entertainment is no longer restricted to those able to afford expensive gaming consoles and top of the line personal computers but rather cuts across all socio-economic statuses. Given these two facts, that is, the immense popularity of video-games, and their potential reach, understanding the benefits and harms associated with video-game play is more important than it ever has been. Indeed, psychologists have long been interested in video-games. The advent of the home computer revolution occasioned the first wave of research into the psychological impact of video-gaming from roughly the late 1970s through to the mid-1980s.[13] Among other things, research into the possibility of improving visual-spatial skills in general, and Mental Rotation in particular, formed part of this early work and was met with some success. Dorval and Pepin (1986), for instance, looked at the effect that 8 sessions (spread over 6 weeks) of training on a 3-Dimensional action video-game (Zaxxon) would have on subjects' performance on a test of 3-D Mental Rotation, namely, the Space Relations Test of the Differential Aptitude Tests. They found that, post-training, subjects who had been assigned to the experimental condition that received video-game time showed significantly higher spatial skill scores than those subjects who had been assigned to a control group. Interestingly, both male and female subjects in the experimental condition seem to have benefited equally from the training (Dorval and Pepin, 1986). Similarly McClurg and Chaille (1987) found that, after receiving training on a number of video-games, both male and female subjects – children from 5th to 9th grade – demonstrated an equal, statistically significant, improvement in performance on The Mental Rotations test compared to a control group. More recently, sophisticated imaging techniques have given psychologists and neuroscientists the ability to directly observe how video-game training changes the brain itself. An interesting example, related to the present study, is the work done by psychologist Richard Haier and colleagues (Haier et al., 2009) in which they demonstrated that 1.5 hours of Tetris play a week over 3 months caused both structural (thickening of the cortex) and functional changes (decreased blood oxygen level dependent responses) to the brains of 28 females (ages 12-15).

In the present section we turn our attention to empirical work that has made use of Tetris as a cognitive training task. The review of this work is presented in two sections. The first section reviews those studies that have used Tetris to investigate factors (primarily sex) that impact MR performance while the second section reviews those studies that focus specifically on the extent to which skills acquired through Tetris training are transferable to non-game contexts.

2.2.1 Tetris, individual differences, and MR performance

In section 2.1.2 we highlighted a small but, for our purposes, important part of the substantial literature investigating those factors that affect MR performance (Voyer et al. 1995; Linn and Petersen 1985). As we have seen, the existing work

[13] Green and Bavelier (2006) provide an extremely broad, but thoroughly readable, introduction to the field and survey a wider range of work than the present literature review.

on MR performance has given significant attention to sex-dependent differences because sex has been the most consistent correlate of MR ability (Linn and Petersen, 1985). It is not surprising then that part of the existing work utilizing Tetris as a MR training task has focused on sex dependent differences in Mental Rotation performance.

In one of the earliest Tetris/MR studies, Okagaki and Frensch (1994) made use of the game in an investigation of sex-differences and the effects of video-game practice on measures of Mental Rotation ability, Perceptual Speed, and Spatial Visualization in older adolescents. For our purposes, it is important to note that two of the measures of visual-spatial performance that they chose to focus on – Mental Rotation and Spatial Visualization[14] – were selected specifically because the researchers believed that they are *required* when playing Tetris (Okagaki and Frensch, 1994). Here we find what is perhaps the first explicit statement in the Tetris/MR literature of the assumption that MR is an essential component in Tetris game play. In their discussion of Tetris the researchers pay almost no attention to the dynamics of the game itself but, rather, merely describe how that game is played and then *assert* that Mental Rotation is required. On the back of this assertion the researchers hypothesize that if Tetris training has any impact on the subjects, it should manifest itself in the measures of Mental Rotation and Spatial Visualization performance and *not* in Perceptual Speed.

In their Experiment 1, Okagaki and Frensch (1994) sought to determine whether there was a sex-dependent difference in, firstly, overall performance in a battery of visual-spatial tests before training, secondly, in Tetris performance, and, finally, in the impact that video-game playing had on visual-spatial performance. Subjects (N=57, 29 female), all undergraduate psychology students (Mean Age = 19.85, SD = 3.52), were administered four paper-and-pencil tests, taken from the French Kit, assessing both 2-D and 3-D Mental Rotation ability, spatial visualization, and perceptual speed. 2-D and 3-D Mental Rotation performance was assessed using the Card Rotations Test (CRT) and Cube Comparison Test (CCT) respectively, both of which are described above from page 12. Perceptual Speed was measured using the “Finding As” task in which subjects are presented with an array of words and are required to cross out all

[14] Linn and Petersen define Spatial Visualization tasks as those that "involve complicated, multistep manipulations of spatially presented information" (Linn and Petersen, 1985). Importantly, Mental Rotation may be one of the *sub-processes* involved in an instance of Spatial Visualization, which is why we should expect certain kinds of Spatial Visualization tasks to be improved if there is an improvement in MR.

words containing the letter "A". Finally, Spatial Visualization was measured using the Form Board Test (FBT) which presents subjects with a target shape and requires them to identify which shapes, from a set of five, could be combined to make the target. In order to measure Tetris performance the researchers recorded the mean number of points and the mean number of lines cleared for the subjects' first and last training sessions.

Subjects were randomly assigned to either an experimental group, which was required to play a total of twelve 30 minute sessions of Tetris, or a no-practice control group. Pre- and post-tests consisted of two different versions of the CRT, CCT, FBT, and "Finding As" task.

Results from the pre-test showed males significantly out-performing females on the CRT ($p < .05$), CCT ($p < .05$), and the FBT ($p < .05$) but not in the test of Perceptual Speed, "Finding As" ($p > .31$). A sex-dependent difference was also found in Tetris performance for the first training session with males achieving a significantly higher number of points ($p < .001$) and number of lines cleared ($p < .001$) than their female counterparts.

After receiving 6 hours of Tetris practice the researchers noted not only a significant improvement in the number of points ($p <.001$) and number of lines ($p <.001$) cleared for both males and females but also that the degree of improvement did not depend on sex – both male and female subjects' *Tetris* performance improved equally with training. What is most important are the patterns of change between pre- and post-tests for the four measures of visual-spatial ability. Only the male subjects in the experimental group showed a significant improvements – on the CCT ($p < .001$) and FBT ($p < .05$) – compared to male control subjects. Okagaki and Frensch's (1994) results thus seemed to demonstrate that Tetris training *is able to* improve certain kinds of visual-spatial abilities, but that this kind of training favours males.

In a more recent study, De Lisi and Wolford (2002) investigated sex-dependent differences in MR ability in children (between 8 and 9 years old) and, like Okagaki and Frensch (1994), were able to demonstrate that Tetris training resulted in improved MR performance. For their pre- and post-tests, subjects (N=47, 23 Female) were administered a version of the French Kit's Card Rotation Task (CRT), the level of difficulty of which had been adjusted in order to be more appropriate for children. Subjects were then allocated to either an experimental group or control group. The experimental group was assigned Tetris as a training task, while the control group was assigned the game "Where in the USA is Carmen Sandiego", an educational game designed to test knowledge of geography and history. Subjects played their assigned games for approximately 330 minutes in 11 sessions spread over the course of a month.

All Tetris scores were recorded allowing for the calculation of two Tetris performance measures by averaging the highest scores of the first (*beginning average score*) and last (*ending average score*) three sessions for each subject. A median-split on pre-test scores divided subjects into groups of high (n=24, 8 female) and low (n=23, 15 female) MR ability. Their results showed the familiar significant sex-dependent difference in MR performance at pre-test with males outperforming females ($p < .05$). As with Okagaki and Frensch's (1994) findings, control and experimental groups showed comparable pre-test MR performance but after the training period the experimental group significantly outperformed ($p < .01$) the control group on post-test MR performance, suggesting that Tetris training had lead to an increase in MR performance. However, contrary to Okagaki and Frensch's (1994) findings, it was the female subjects who benefited most from the training – with the sex-dependent difference that was evident in the pre-test scores practically eliminated in the experimental group at the end of their training period.

Interestingly, in their analysis of the Tetris performance of the experimental group, De Lisi and Wolford (2002) found that although subjects' pre-test MR scores were not correlated to their *beginning average score* in Tetris ($p > .05$), their post-test MR scores *were* significantly correlated to their *ending average score* ($p < .05$). The researchers suggest that what might explain the post-training correlation of Tetris performance and MR scores is a shift in the subjects' Tetris strategy. Specifically, the correlation could be explained if at the end of training subjects were relying *more* on Mental Rotation while playing Tetris than they were in their first three sessions. The researchers unfortunately leave this tantalizing possibility unexplored.

How can we explain the differences in De Lisi and Wolford's (2002) and Okagaki and French's (1994) findings? It's possible that the most important difference was the average age of the participants. As we've already seen with both Kail and Park's (1990) work on MR training, as well as the pair of studies by Weidenbauer and colleagues (Wiedenbauer et al., 2007; Wiedenbauer and Jansen-Osmann, 2008), children and adults respond differently to MR training with children tending to show greater gains in MR performance over similar training periods. This developmental explanation doesn't seem to hold though, as a 2008 study by Cherney focusing on subjects of roughly the same age (Mean Age = 19.1, SD = 1.4) as Okagaki and Frensch's (1994) demonstrated patterns of improvement in MR performance similar to those found in De Lisi and Wolford (2002). In addition to performance differences across sexes, Cherney investigated a number of further factors in both the *kind* of training subjects receive as well as other potential individual differences between subjects that have been hypothesised to have an effect on MR performance. With regards to the kind of training the subjects receive, Cherney (2008) sought to address the possibility that distributed practice – practice sessions spread out over time – of the Tetris training task might be more effective than massed practice as there is evidence suggesting that distributed practice of simple motor tasks produce better results and longer retention periods of the skills acquired than does massed practice (Donovan and Radosevich, 1999). With regards to individual differences, Cherney was interested in two main factors that could affect MR performance.

Firstly, Cherney was interested in assessing whether anxiety may have a negative affect on female MR performance as anxiety levels have been shown to negatively affect performance on a number of tests of cognitive ability (memory, analogical reasoning, etc.) (Cherney, 2008). Furthermore, females have been found to demonstrate higher levels of test and mathematics anxiety than males, Cherney hypothesised that this could be a contributing factor to the sex-dependent differences almost always demonstrated in tests of MR performance (Cherney, 2008). Secondly, she wanted to assess the impact that previous spatial experiences and practice in quantitative reasoning (especially in mathematics and science) would have on overall performance as well as sex-dependent differences in MR performance (see, for example, Voyer and Sullivan, 2003 and Ozel et al., 2004).

In order to address these issues, Cherney randomly assigned 61 undergraduate students (31 females) to either one of two experimental groups, or a control group. The two experimental groups were assigned the task of playing either a simple 3-Dimensional driving game (Antz) or Tetrus, a generic Tetris clone. The control group was assigned paper and pencil word puzzles. Pre- and post-test measures of MR performance were the 2-D Card Rotation Test (CRT) and the Vandenberg-Kuse Mental Rotation Test (VMRT), both of which are described above. To assess test anxiety levels, all subjects were administered the State-Trait Anxiety Inventory (STAI). In addition subjects were asked to complete a questionnaire regarding previous video-game experience, handedness, and their background in mathematics, science, and sports. Finally, subjects were required to complete a short mathematics test consisting of six questions of varying difficulty.

Roughly half of each of the three conditions were administered their training in four hour long sessions in a single week (*massed training*) while the remainder were administered the same amount of training over two weeks (*distributed training*).

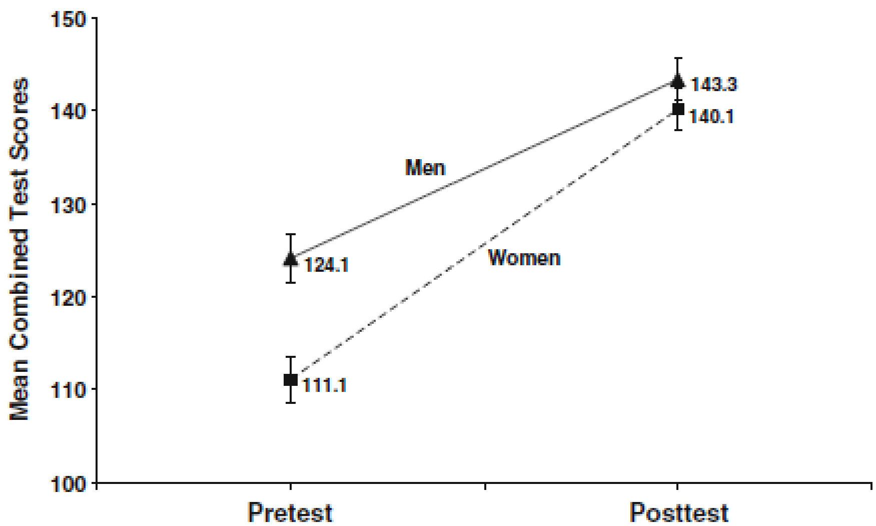

Figure 5: Means and standard error for pre-test and post-test scores on VMRT and CRT (Cherney, 2008)

Their results showed, like both of the earlier Tetris studies, that the experimental groups' MR performance improved significantly in comparison to the control group but that – as with De Lisi and Wolford (2002) – women's gains were, overall, significantly greater than men's ($p < .05$). A regression analysis run on the groups from a median split on mathematical achievement further revealed that high mathematical performance, sex, and type of practice condition significantly predicted improvement. Interestingly, the 3-D driving game proved to be *slightly* more effective for training MR than Tetris on the VMRT – possibly because the VMRT requires rotation in 3-Dimensions. Finally, Cherney found that massed training improved MR performance significantly more ($p < .05$) than distributed training.

What is particularly important about Cherney's results for the present study is that they demonstrate that it is possible for Tetris training to cause measurable improvements in MR performance with only 4 hours of training.

2.2.2 Tetris training and Transfer

The most important and thorough work that has been done on the transfer effects of Tetris training is the Doctoral work of Valerie Sims, presented in (Sims and Mayer, 2002). In this study, Sims and Mayer were interested in determining the extent to which those skills learned and practised during Tetris play transferred to other measures of visual-spatial ability. In order to classify the extent of any transfer effects they observe, Sims and Meyer delineate three possible outcomes corresponding to views about the nature of visual-spatial skills. These are, firstly, *transfer of general skills*, secondly, *transfer of specific skills* and, thirdly, *transfer of specific skills in context*. The *transfer of general skills* view takes visual-spatial ability as a more or less unified faculty that can potentially be altered and improved as a whole through any kind of visual-spatial training. This view predicts that Tetris training should cause improvements in all or most measures of visual-spatial ability. The *transfer of specific skills* view takes visual-spatial ability as being comprised of a number of disparate abilities, of which Mental Rotation is one. This view predicts that training using a task that ostensibly engages Mental Rotation should improve *only* Mental Rotation, and that this improvement should be evidenced across all contexts in which this ability is used. The final view of transfer Sims and Mayer describe is *transfer of specific skills in context*. According to this view visual-spatial ability is not only comprised of several disparate abilities but also that experience with video-game play serves only to "improve component skills using the same mental representations as are required in the game" (Sims and Mayer, 2002: 99). This view predicts that if a subject is subjected to Tetris training she should only demonstrate improvements in MR performance in those instances where the target and comparison images are the same as those used in the game itself – this view is consistent with the notion that improvement in Metal Rotation is primarily *instance based* (see section 2.1.4 above).

Sims and Mayer's study comprised two parts. The first part consisted of measuring the differences between expert Tetris players (N=53, 17 females) and non-video-game players (N=45, 26 females) on a range of visual-spatial abilities. The second part consisted of a longitudinal study in which 16 female subjects were assigned to either an experimental group that was required to practice Tetris for 12 hours or a no-practice control group. Group assignment ensured approximately equivalent performance on pre-test measures of visual-spatial ability. In this part of the study the researchers sought to determine whether there exists a causal relationship between Tetris training and visual-spatial performance. For both parts of the experiment Sims and Mayer used a range of measures that would represent nearer or farther transfer of Tetris expertise. These were, in descending order of nearness of transfer, four Shepard-Metzler style MR tests (using Tetris shapes, Non-Tetris shapes, Tetris-like letters, and Non-Tetris-like letters as stimuli), two computerized Form Board Tests (one using Tetris and the other non-Tetris shapes), a Card Rotations Test, a paper and pencil Form Board Test from the Kit of Factor Referenced Cognitive Tests, and a Paper Folding Task[15], also taken from the Kit of Factor Referenced Cognitive Tests.

In the first part of the study these measures were only administered once to both expert and novice Tetris players while in the second part of the study these measures were used as pre- and post-tests.

When comparing expert and novice performance, Sims and Mayer found that the only statistically significant difference in performance between the two groups was to be found in the Mental Rotation of Tetris shapes ($p < .01$) and non-Tetris shapes ($p < .05$), a finding that seemed to be consistent with the notion that Tetris training yields only transfer of specific skills in context. However, the results of the longitudinal study revealed that 12 hours of Tetris training led to no statistically significant differences between experimental and control conditions on post-test measures of spatial ability.

In contrast to these negative results, a more recent longitudinal study by Boot et al. (2008) found that subjects demonstrated transfer effects after approximately 21 hours of Tetris training. As part of a larger study investigating the effects of different kinds of video-games on several measures of visual-spatial ability, attention, reasoning, and executive control Boot et al. (2008) administered a Shepard-Metzler style MR test – whose target and comparison images were based on Tetris shapes – to 7 groups of students (6 experimental groups, 1 control group) as part of a battery of cognitive tests. Subjects were then required to play 21 hours of an assigned game over several weeks, after which they were administered the same battery of cognitive tests. It was found that after the training period subjects who had been assigned Tetris as their cognitive task had improved significantly more ($p < .05$) than those who had been

[15] In this task subjects are presented with a series of images depicting a number of steps in which a piece of paper is being folded. The final step shows a hole punched through the folded paper. Subjects are then shown five images that possibly depict the punched paper when it is unfolded, only one of which is correct. The subject is required to determine which of the five unfolded images is correct.

assigned other kinds of video-games or those assigned to a no-practice control group. Importantly, the Tetris group showed no other significant differential improvements for any of the other measures when compared to subjects assigned to the other groups.

It is possible to see (Boot et al., 2008) as being consistent with Sim and Mayers' (2002) findings. Although Sims and Mayers failed to find any evidence of transfer with 12 hours of Tetris training, one might argue that that transfer effects *would have* emerged if subjects were given a longer training period. Further, Boot et al. (2008) saw their Tetris group demonstrating *transfer of specific skills in context* – that is, Tetris training seemed to only improve MR performance and not any of the other visual-spatial measures of interest – much like the results of part one of (Sims and Mayer, 2002).

What is most significant for the present study about Boot et al. (2008) is that in their discussion of their results they raise the question of what skills are *actually* being exercised by Tetris players, unlike any of the other studies reviewed here. Specifically, they make mention of Kirsh and Maglio's (1994) work on Tetris that suggests that expert Tetris players may make use of epistemic actions to offload mental computations (such as Mental Rotation) on to their environment – we review this work in detail in the next section. Boot et al. (2008) also go on to suggest that future studies should investigate these kinds of strategies that are learned while playing video-games as they may represent an important, and neglected, aspect of video-game expertise effects. The present study was undertaken partly as a response to this suggestion.

2.3 Epistemic Actions in Tetris

2.3.1 Epistemic versus Pragmatic actions

In their seminal 1994 paper, Kirsh and Maglio introduce the distinction between PRAGMATIC ACTIONS and EPISTEMIC ACTIONS. Pragmatic actions are understood as those physical actions that serve to bring an agent closer to some goal state.[16] For example, if you wish to have a tidy living room – your goal state – then the act of moving a pair of running shoes from the room to your bedroom cupboard is a clear example of a pragmatic action. By moving the shoes you have advanced yourself one step closer to your goal state of a tidy room.

[16] In the present study we mainly consider human agents. However, the distinction between Pragmatic and Epistemic actions is potentially applicable to the actions of any kind of information processing and goal seeking agent that is able to sense the state of, and physically affect, its environment.

In contrast, an epistemic action does not *primarily* serve to bring the agent closer to some goal state but rather it is to be understood as a physical action intended to change the agent's *informational state.* More precisely, an epistemic action is defined as being a physical action that either (1) relieves memory requirements for mental computation, (2) reduces the number of steps required for mental computation, or (3) reduces the probability of errors that might occur during mental computation (Kirsh and Maglio, 1994). For instance, you may take a pair of running shoes from your cupboard and place them at the threshold of your front door in order to remind yourself that you arranged to go running with a friend. By placing the shoes at the front door you are effectively using your environment as a substitute for, or compliment to, your natural powers of recall.

It's important to recognise that the two kinds of actions – pragmatic and epistemic – are not mutually exclusive. It is possible for a physical act to both bring one closer to a goal state while simultaneously affecting the acting agent's informational state in accordance with Kirsh and Maglio's definition of epistemic actions. To return to our running shoe example, leaving your shoes near the front door may *both* play the epistemic role of reminding you that you've arranged to run *as well as* playing the pragmatic role of actually getting ready to run by placing your shoes in a convenient location.

2.3.2 An argument for the existence of Epistemic actions in Tetris

Kirsh and Maglio themselves point out that the notion that physical actions can serve to make cognition easier, faster, or more reliable has long been established – in the decade preceding Kirsh and Maglio's work there were, for example, studies published detailing the ways in which office workers arrange their desks to remind them to perform certain tasks (i.e. an action that reduces memory requirements) (Malone, 1983) and studies on the cognitive benefits of making external representations of one's ideas, such as writing equations on a blackboard or sketching diagrams (Riesberg, 1987).

What Kirsh and Maglio sought to demonstrate in their 1994 paper is that actions that serve to simplify mental computation are far more pervasive than cognitive psychologists had previously recognised. Using data collected from subjects playing Tetris they attempted to show an example of the existence of epistemic actions in a task that is rather different from those that had previously been investigated. Firstly, Tetris is unlike, say, the process of organizing an office desk in being extremely fast paced. Secondly, it is not a *symbolic* process in the same way that externalising one's thoughts on paper as equations or sketches might be. If Kirsh and Maglio have been successful in establishing the existence of epistemic actions in this fast paced, non-symbolic task we then have a, at least *prima facie,* reason for thinking that epistemic actions may exist in a whole range of previously unexamined activities that may not be obvious candidates for supporting actions that compliment cognitive processes.

An important part of Kirsh and Maglio's work is their challenge to theories of action and planning that don't recognize the existence of epistemic actions. In the present study we are not particularly concerned with this aspect of their work – we are ultimately concerned with what kinds of epistemic actions are available to Tetris players and what impact this has on the game's efficacy as an MR training tool. However, Kirsh and Maglio's *challenge* to classical theories is the first step in their argument for the existence of epistemic actions in Tetris (Kirsh and Maglio, 1994) and so an understanding of it is essential.

Very roughly, any theory failing to recognise the existence of epistemic actions in behaviour that *actually contains* epistemic actions might have a problem in determining whether the observed behaviour is *optimal.* Below we offer a reconstruction of an argument found in (Kirsh and Maglio, 1994) predicting what we should expect optimal Tetris play – specifically, the expected pattern of zoid rotation – to look like from a theory of action that fails to recognise epistemic actions. Note, it is only by contrasting the *actual* game play data of expert Tetris players to these kinds of predictions that Kirsh and Maglio find the room to interpret some in-game moves as epistemic actions.

Premise 1: Expert Tetris players will consistently use close to the minimum number of moves to place a zoid.

During a Tetris episode, the *shortest* path from the beginning state (when the zoid enters the playing area) to the goal state (the final placement of the falling zoid at the bottom of the well) will be the path that contains the *fewest* number of rotations and translations. Any theory that fails to recognise the existence of epistemic actions in Tetris is almost certainly going to equate the *shortest* path with the *optimal* path. In a time limited game any extraneous moves will be interpreted as, either, a waste of effort (i.e. they will need to be undone, they take unnecessary time, etc.) or simply errors that will need to be corrected. Note that this premise rests on the additional assumption that part of what it means for an individual to be an expert Tetris player is that their behaviour tends towards optimality – this is an assumption that is shared by Kirsh and Maglio but, as we shall see, they disagree over just *what* is to be counted as optimal.

Premise 2: If a Tetris player consistently uses the minimum number of moves required to place zoids, then over a large number of episodes the number of rotations per zoid type should average half of the number of rotations that can be performed before the zoid is back in its original orientation.

When a zoid emerges from the top of the screen in Kirsh and Maglio's version of Tetris, it emerges at a random orientation. If we further assume that, on average, a zoid will be placed in any of its orientations with equal probability, then – if a Tetris player is using the minimum number of moves required to place their zoids – over a large number of episodes the average number of rotations per zoid type should be roughly half the total number of rotations that can be performed before the zoid is back in the orientation in which it emerged.

For example, the T-Shaped zoid can make three 90°rotations before it returns to its original orientation. We should then expect that in the long run the average number of rotations for this zoid-type will be 1.5.

Conclusion: Over a large number of episodes Expert Tetris players' average number of rotations per zoid type should be half the number of rotations that can be performed before a zoid is back in its original orientation.

Note that this argument doesn't *guarantee* that Expert Tetris players will display the predicted pattern, but it does make a clear prediction about what we should reasonably expect regarding patterns of zoid rotations if no epistemic actions are used *and* the player is using the minimum number of move to place their zoids. This is crucial because it helps us establish a base *against which* we're able to judge whether a player is rotating, on average, more or less than expected.

According to Kirsh and Maglio's results (Kirsh and Maglio, 1994; Maglio, 1995; Maglio and Kirsh, 1996) expert Tetris players rotate their zoids more than is predicted by classical theories of action and planning. Table 7, on page 64, shows a detailed comparison of the average number of rotations Kirsh and

Maglio's subjects made per-zoid type compared to the average number predicted by the argument given above.

In accounting for the existence of these extra rotations and translations, Kirsh and Maglio challenge the argument's first premise. Specifically, if there exist strategies in Tetris that reduce the need for mental computation – that is, epistemic actions – it might *not* be the case that expert Tetris players will use the minimum number of moves in placing their zoids. Optimal *human* Tetris play might not be equivalent to placing zoids in the minimum amount of moves but could include use of extra movements that serve to make Tetris cognition more efficient.

Of course, Kirsh and Maglio can't merely assert that the extra movements in their data *just are* epistemic actions – the patterns of over-rotation and over-translation generated by their expert Tetris players might just as easily have been generated by, say, simple mistakes or changes of mind mid-placement (Destefano et al., 2011), and no doubt some of the extra movements actually *are* just mistakes or changes of mind. Kirsh and Maglio's argument for assigning at least some of the extra movements in their data an epistemic rationale has two parts. The first part consists in providing an account of possible epistemic uses of rotation and translation in Tetris that are compatible with their players' actual behaviour. The second part consists in providing some evidence that expert Tetris players actually make use of epistemic actions. They attempt to do this by providing evidence suggesting that Tetris players actually make more "extra" movements as they become more skilled – presumably because as they improve they make greater use of epistemic actions. We detail both parts of their argument below.

Kirsh and Maglio's data were drawn primarily from two experiments that formed part of Maglio's PhD research (Maglio, 1995). In the first of these experiments, Kirsh and Maglio had subjects ($N = 33$, 6 female) between the ages of 19 to 32, play six games of Tetris in a single session. The version of Tetris used in these sessions collected all subjects' keystrokes and accurate timing information enabling the researchers to recreate the full detail of the games, allowing them to analyse their subjects' game play in depth. The 33 subjects were partitioned into three groups – 11 beginners, 12 intermediates, and 10 experts – using k-means clustering on the means and standard deviations of their Tetris scores (Maglio, 1995). The second experiment was a longitudinal study in which they collected twenty hours of in-game Tetris data for two subjects, the details of this experiment are given below in Section 2.3.5.

2.3.3 Epistemic uses of Rotation

In this section we discuss a number of potential epistemic uses of physically rotating zoids (Kirsh and Maglio, 1994).[17]

Rotating early to unearth new information: In the version of the Tetris Kirsh and Maglio used to train their subjects, zoids do not appear in the playing area all at once but are initially only partly visible and are revealed gradually as they descend from the top of the screen. In many cases the initial image produced by the partially emerged zoid is ambiguous in the sense that it could be produced by more than one zoid type. These partially emerged zoids can either be ambiguous in shape only or in both shape and position. Both kinds of ambiguity are illustrated in Figure 6.

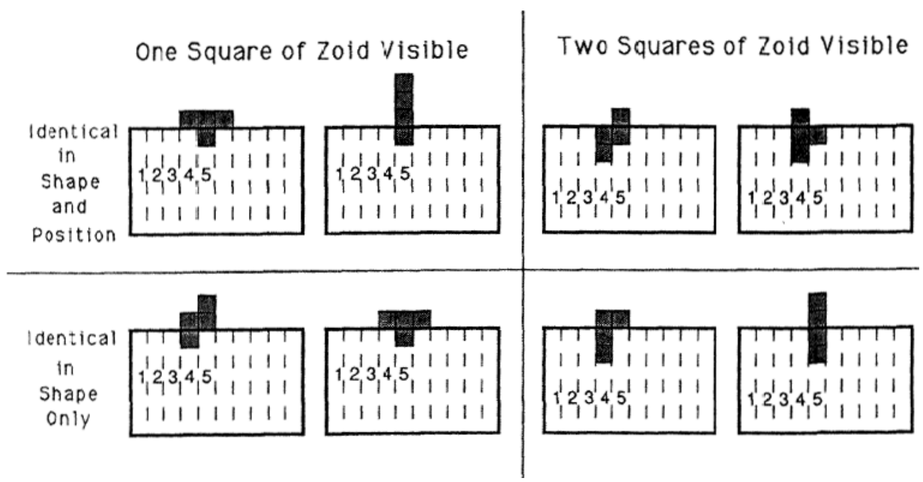

Figure 6: When zoids emerge from the top of the screen they are only partly visible in some cases making their identity ambiguous (Kirsh and Maglio, 1994)

Kirsh and Maglio state that the rate of descent for their version of Tetris was one square per every 150ms. This meant that, depending on the zoid's initial orientation, a Tetris player might need to wait up to 450 ms to properly identify a zoid's type. What their data reveal is that Tetris players sometimes rotate their zoids *before* they are fully emerged (Kirsh and Maglio, 1994). Specifically, they

[17]The discussion of the epistemic uses of rotation in Kirsh and Maglio's "On distinguishing Epistemic from Pragmatic Action" (1994) is often framed in terms of their own, particular, model of Tetris cognition. As their models are only peripherally related to the present study we have opted to present their work on epistemic actions without reference to this part of their work.

found that subjects were prone to rotating zoids before they were fully visible if their images were ambiguous in shape. Further, subjects were even *more* likely to perform this routine if the partly visible images were also ambiguous in position (Kirsh and Maglio, 1994). Kirsh and Maglio argue that the best interpretation of these early rotations is that subjects are rotating in order to disambiguate and identify the zoid. By rotating the zoid while it is still partially hidden subjects see parts of the zoid at different orientations allowing them to infer its actual shape up to 300 ms earlier than would have been possible if they had let the zoid emerge naturally.[18]

Rotating to save Mental Rotation Effort: For the present study the most important proposal regarding the epistemic use of rotation is that expert Tetris players may rotate their zoids to save Mental Rotation effort. Kirsh and Maglio suggest the when Tetris players are faced with the task of matching a zoid to the contour of the bottom of the playing area they have two options for testing the zoid's goodness-of-fit at all possible orientations. Firstly, the Tetris players could use MR to reorient the zoid and then match the resultant "mental image" against the contour. Secondly, Tetris players could rotate the zoid physically and then – after whatever process of visual encoding is required for the player to perceive the zoid at its new orientation – match the physical image itself to the contour.

Kirsh and Maglio present data suggesting that Tetris players might prefer the second option because it's simply *faster* for them to physically rotate the zoid 90°by pressing a key than it is for them to perform the same operation using MR. In a small experiment ($N = 3$) Kirsh and Maglio used a Shepard-Metzler style MR test to estimate the average time it takes for Tetris players to Mentally Rotate Tetris shapes. Their estimation for the fastest MR performance was around 800-1200 ms per 90° rotation while they estimate that Tetris players are able to physically rotate zoids at 100-400 ms per 90° (Kirsh and Maglio, 1994). Other than simply being faster than Mental Rotation, physical rotation has the added benefit of practically eliminating any costs, in terms of memory

[18]Interestingly Destefano et al. (2011) argue that according to Kirsh and Maglio's own definition, rotating to unearth the zoid's type doesn't strictly qualify as an epistemic action. Specifically, they point out that these early rotations do not help improve cognition by reducing the memory required for related mental computations, reducing the number of steps involved in related mental computations, or reducing the probability of error in mental computation. Early rotation uncovers previously unavailable information that is a prerequisite for planning, rather than being a physical offset of a process that could have taken place through mental computation alone.

or attentional resources, that might be associated with sustaining the mental image being rotated (Kirsh and Maglio, 1994).

Rotating to Facilitate retrieval of zoids from memory or identifying a zoid's type: Kirsh and Maglio (1994) also raised the idea that seeing a zoid at multiple orientations may serve to speed up certain cognitive processes, suggesting that some of the extra rotations observed in their data might be evidence of subjects priming their own perception or recall. While the suggestion is only briefly sketched in the earlier work, Maglio – with Wenger and Copeland – recently revisited this hypothesis in a series of experiments investigating how rotating zoids may act as primes for Tetris players (Maglio et al., 2008). In their experiments 1 through 3 Maglio and his colleagues investigated priming effects that improved subjects' response times in a task requiring them to indicate whether a zoid would fit the contour of a board. The experiments consisted of a series of trials each of which had subjects observe a series of *preview zoids*, each displayed on their own for 250ms, before being shown a *final test zoid* set above a contour reminiscent of the contour at the bottom of a Tetris playing area. Subjects were then required to indicate, as quickly and as accurately as possible, whether or not the final test zoid matched, or would fit into, the contour below it.

In their experiments Maglio et al. were able to control whether the final test zoid appeared in the series of preview zoids, where it appeared in the series, how many times it appeared, and whether – if a zoid appeared more than once – it appeared at multiple orientations. Their results showed that RTs were faster if the final test zoid appeared in the preview series than if it didn't. Further, RTs were even faster if the final test zoid appeared multiple times at multiple orientations in the series. Finally, RTs were also shown to improve if the previews of the final test zoids were shown earlier in the preview series than if they were nearer to the end. These results (Maglio et al., 2008) provide empirical support for Kirsh and Maglio's initial suggestion – that is, it may be the case that physically rotating a zoid early in an episode may prime Tetris players' recognition and/or recall.

2.3.4 Epistemic uses of Translation

Kirsh and Maglio (1994) identify one clear epistemic use of translation. They note that in about 1% of the cases where their subjects choose to drop their zoid manually, the command to drop the zoid is immediately preceded with a translation routine in which subjects quickly move their zoids to the nearest wall and then back to the final column in which they drop it. The data also reveals that as the distance to drop the zoid increases, so does the likelihood that the subjects will perform this routine (Kirsh and Maglio, 1994). Kirsh and Maglio suggest that the point of this translate-to-wall-and-back routine is to verify that they zoid is being dropped into the correct column. Subjects achieve this by counting the number of columns there are between the wall and their intended "drop zone" and then match this count with an equal number of zoid translations from the wall (Kirsh and Maglio, 1994).

2.3.5 Do Epistemic Actions increase with skill?

As mentioned above, part of Maglio and Kirsh's work on epistemic actions consisted of a longitudinal study investigating whether, and how, the average number of rotations and translations Tetris players make would change as their skills improved with practice (Maglio and Kirsh, 1996). They point out that traditional models of skill acquisition predict that, with practice, individuals should be observed making fewer mistakes, optimising their movements, and generally speeding up (Maglio and Kirsh, 1996). These models would predict that, in the case of Tetris, as individuals become more experienced we should expect to see their movements get faster, as well as the number of extraneous actions they make decrease. The latter can be expected as experienced players should make fewer mistakes, eliminating those moves needed to correct them.

In order to test this hypothesis Maglio and Kirsh had subjects (N=2), who had no previous experience with the game , play 20 hours of Tetris in their laboratory. As in their earlier studies (Kirsh and Maglio, 1994), all in-game data were recorded for analysis.

Kirsh and Maglio's subjects did show an overall increase in speed, as predicted by traditional models (Maglio and Kirsh, 1996). More importantly they also found that the average number of movements that their subjects performed per episode actually increased with practice. This was demonstrated by grouping their subjects' data into three consecutive 6-hour intervals. The mean number of extra rotations per-game were calculated and then averaged for each of the three intervals. Comparing these averages revealed a statistically significant increase ($p < .01$) in the number of extra rotations with time (Figure 7 shows the breakdown of the extra rotations per zoid type at each one of the inter-

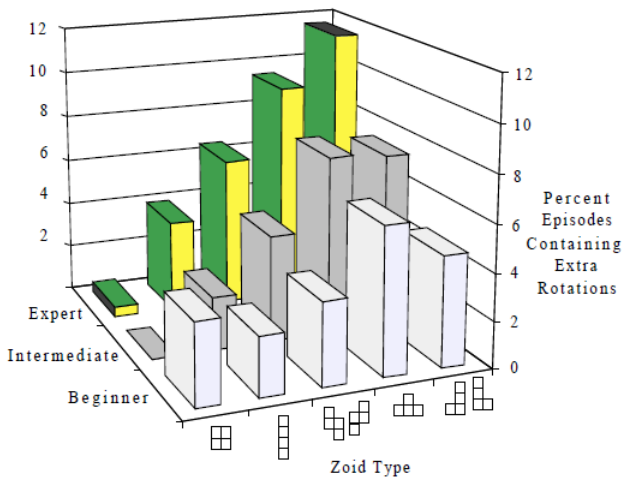

Figure 7: Increase in extraneous rotations at different skill levels (Maglio and Kirsh, 1996)

vals). This increase in extra moves with practice can be neatly explained by the hypothesis that expert Tetris players make use of epistemic actions to aid Tetris cognition. As novice Tetris players gain more experience with the game they might be expected to uncover at least some of these epistemic actions and, with practice, learn how to integrate them into their style of play. If epistemic actions make Tetris players more effective, we should expect their use of them to increase with time and practice.

Destefano et al. (2011) have recently challenged Maglio and Kirsh's assertion that epistemic actions increase with skill with data showing that this may only be the case for early stages of Tetris skill acquisition. In their study they don't deny the existence of epistemic actions but, rather, raise a number of challenges to Kirsh and Maglio's work.

Specifically, Destefano et al. (2011) take issue with the range of expertise represented by Kirsh and Maglio's subjects. For their own study Destefano et al. attempted to get a wide range of Tetris expertise by recruiting subjects at a convention for fans of Science Fiction, Fantasy literature, Japanese Anime, and video games – a venue where one can reasonably expect to find a wide range of video-game skills.

The first phase of their experiment was presented as a competition – each subject would compete in a qualifying round by playing two games of Tetris, their qualifying score being the highest from their two games. As with Kirsh and Maglio's experiments, Destefano et al.'s version of Tetris recorded all game data and player moves. Once qualifying rounds were over, the top eight competitors played a series of one-on-one elimination matches, a process that ultimately left a single overall winner. For the second phase of the experiment the top three contestants were invited back to the researchers' laboratory where they played as many rounds of Tetris as they could in one hour. Using this process of recruitment the researchers managed to address a further issue with Maglio and Kirsh's 1996 study, namely, the fact that their sample consisted of a mere two subjects – Destefano et al. (2011), on the other hand, managed to collect game data from 59 subjects during their competition.

Ultimately, Destefano et al.'s subjects represented a wide range of skill levels with Tetris scores in the qualifying round ranging from as low as 867 points to as high as 236,305 points (Destefano et al., 2011). These subjects were then placed into 5 distinct groups of increasing Tetris skill, determined by their highest scores. Interestingly, Destefano et al. (2011) try to infer where the subjects from (Maglio and Kirsh, 1996) would be placed in their 5 groups and their best estimate is that Kirsh and Maglio's subjects, who played 20 hours of Tetris, would be located in either their first or second level of Tetris skill. If this is the case – and Destefano et al. (2011) are clear that their inference is inexact – then the range of skills represented in (Destefano et al., 2011) is indeed much wider than in Kirsh and Maglio's work.

Their results show that if one examines the incidence of epistemic actions across the 5 skill levels represented by their subjects, one sees an initial increase in the use of epistemic actions between the first and second skill levels, after which the incidence of epistemic actions drops as the subjects become more skilled. It's interesting to note that their most skilled player's highest scoring game of Tetris contained only a *single* instance of over-rotation (i.e. possible rotation based epistemic action) in 1281 episodes, as well as the fewest number of translation based epistemic actions compared to any other game in their dataset (Destefano et al., 2011).

3 Aims and Rationale

The fact that Tetris requires players to physically rotate zoids under time limited conditions has served to mark it out as a task for investigating the ways in which video-game play affects MR performance. Underlying this identification of Tetris as a MR training task is the assumption that Mental Rotation is *required* or, at the very least, *exercised* when humans play Tetris (Okagaki and Frensch, 1994; De Lisi and Wolford, 2002; Sims and Mayer, 2002). Further, this assumption seems to be somewhat justified as we have seen that there is in fact evidence to suggest a causative link between playing Tetris and improvement in Mental Rotation ability (Okagaki and Frensch, 1994; De Lisi and Wolford, 2002; Sims and Mayer, 2002). However, Kirsh and Maglio's work (Kirsh and Maglio, 1994; Maglio and Kirsh, 1996) suggests that Tetris players can, and sometimes do, forgo purely Mental Rotation in favour of rotating their zoids in physical space. If this is the case then it is possible that previous studies of Tetris and MR performance have misconstrued the nature of the cognitive task posed by Tetris by overemphasizing the role played by Mental Rotation. The central issue that the present study was designed to address is whether the effectiveness of Tetris as a training tool for Mental Rotation is affected by the fact that there exists a class of actions that reduce, or potentially eliminate, the need to engage in Mental Rotation while playing the game.

To address this central issue we set out to answer the following question:

Research Question 1: If a group of participants are trained using a version of Tetris modified in such a way that rotation based epistemic actions are made difficult or impossible, is there a measurable difference in post-test MR performance when compared to a group trained using a standard version of the game?

In order to try establish whether our subjects who were trained using a standard version of Tetris were making use of epistemic actions we focused on whether they showed an increase in the average number of rotations they made as their Tetris skills improved.

As Destefano et al. (2011) point out, it is often difficult to unambiguously classify extraneous movements in Tetris *as* epistemic actions rather than, say, straightforward errors or as instances of players changing their plans mid-episode. If we assume that players will make fewer mistakes as their skills improve then an increase in average number of rotations does provide a *prima* *facie* reason for ruling out the possibility over-rotations we observe in our skilled Tetris players' games are the result of simple error. Observing an increase in the average number of rotations with increasing Tetris skill would therefore go some way in supporting an interpretation of extraneous movements *as* epistemic actions (Maglio and Kirsh, 1996), although it would not rule out the possibility that skilled players are simply more prone to mid-episode plan changes.

Our second research question was:

Research Question 2: Do subjects trained using a standard version of Tetris show a measurable increase in the average number of rotations they make as their Tetris skill increases?

Finally, in a much more exploratory vein, we were interested in investigating whether there were any correlations between the average number of rotations and MR performance or between average number of rotations and Tetris performance. Although we had no comprehensive expectations, this final research question was motivated by two points that served as supplements to research questions 1 and 2. Firstly, if Kirsh and Maglio are correct that some Tetris players substitute Mental Rotation with the physical Rotation of zoids, it seems to follow that those who do may be receiving *less MR practice* than those players who don't rely on epistemic actions. We were interested to see if this would be reflected in any correlations between the average number of rotations made by subjects assigned to play a Standard version of Tetris and their MR performance on a Shepard-Metzler style test.

Secondly, if Kirsh and Maglio's assertion that the use of epistemic actions *increases* with Tetris expertise is correct, then it is possible that this fact may be reflected in the relationship between Tetris players' average number of rotations and their performance in Tetris.

Our third research question was, then:

Research Question 3: Are there any correlations between the average number of rotations in Tetris and performance in either (a) the pre- and post-tests of MR performance or (b) overall performance in Tetris itself?

4 Methodology

4.1 Research Design

In addressing our first research question a pre-post-control design was employed. Subjects were allocated to either a CONTROL group or one of two experimental groups that we designate the STANDARD and MODIFIED groups.

All subjects allocated to the experimental groups were exposed to at least five hours of Tetris training on one of two versions of Tetris. The Modified group's version of Tetris was designed to restrict the subjects' use of epistemic actions involving over-rotation of Zoids while the Standard group's version of Tetris had no such restrictions. The Control group was assigned a task that did not require Mental Rotation. All assigned tasks are described in detail below. Pre- and post-tests consisted of Shepard-Metzler style Mental Rotation tests, also described below.

Two analyses were performed on the data, the first of which was a one-way Analysis of Covariance (ANCOVA) with group allocation as the independent variable, post-test performance on the test of Mental Rotation ability as the dependent variable, and pre-test MR performance as a covariate. ANCOVA allows for the exploration of differences between groups while simultaneously statistically controlling for an additional continuous variable (Pallant, 2011). In the present study we used pre-test scores as the covariate in order to partial out subjects' pre-intervention MR performance which, as we have seen, can vary widely depending on factors such as the individual's sex (Linn and Petersen, 1985), age (Dror et al., 2005), and socioeconomic status (Levine et al., 2005). While there is some controversy surrounding the use of ANCOVA with pre-existing / intact groups to "control for" pre-existing differences within groups it was initially taken to be unproblematic for the present study given the way in which group assignment was undertaken (described below) (Dimitrov and Rumrill, 2003). However, as the results section below shows, our Control group's pre-test MR scores were substantially better than the Standard and Modified groups' scores which raised the possibility of a bias in our group allocation process. In order to address any potential concerns about the ANCOVA and our group allocation, as well as to present a more comprehensive analysis of the data, we undertook a second analysis, namely, a one-way Analysis of Variance (ANOVA) with group allocation as the independent variable and pre- and post-test difference scores (post-test score - pre-test score = difference score) as the dependent variable.

In order to address the secondary research questions (question 2 and 3 above) a number of measures were drawn from the Standard group's in-game data. For each subject's first and last ten games of Tetris we calculated average high scores and the average number of rotations per zoid type.

For research question 2 we ran a series of paired sample t-tests comparing the average number of rotations, per zoid type, for our Standard group's first and last ten games, while for research question 3 we ran a comprehensive series of tests for correlations between the average number of rotations per zoid type, Tetris scores, and MR performance scores.

4.2 Ethical aspects

All subjects received an informed consent form (see Appendix B) which they were required to sign prior to their participation in the present research. The consent form described the purpose of the study as well as providing details about the nature of the tasks (MR tests and video game playing) involved and overall time commitment associated with participation. Subjects were also informed that by signing the consent form they were granting the researcher permission to access biographical and registration information from the university's computer systems. Finally, subjects were guaranteed that their participation and personal details would remain confidential and that they were free to withdraw from the study at any time, and for any reason, without penalty.

The present research was approved by the University of KwaZulu-Natal's Research Ethics Committee.

4.3 Sample

4.3.1 Recruitment

All participants were students registered at the University of KwaZulu-Natal. Convenience sampling was used, with potential subjects being invited to participate in the research by means of:

1. A3 Posters advertising the research displayed around UKZN's Howard College Campus in areas of high traffic and visibility. See Appendix C for the text of the advertisement.

2. An electronic mail sent to all undergraduate philosophy students requesting participation in the research. This same message was posted on the University's electronic "classified" billboard. Again, the text in Appendix C was used for these advertisements.

3. Advertising the research before lectures. Several well attended undergraduate lectures were identified and, after receiving permission, the researcher presented students with the opportunity to participate in the present research. Students were given an opportunity to ask questions, and those who were interested in participating were directed to informational posters placed outside their lecture halls or the project's website for more information.

All potential participants were directed to the, now defunct, website http://gamesforscience.co.za where they were required to fill in an online survey / sign-up form that recorded the following details:

1. Full name
2. UKZN student number
3. Email address
4. Cellphone number (optional)
5. Age
6. Sex
7. How many hours a week spent playing video-games. Potential subjects were required to select one of the following options – "Less than one hour", "About one or two hours", or "More than three hours"
8. Whether, in the last year, they had played more than five hours of Tetris
9. The day of the week that they were able to participate in the research. Potential subjects were required to select either Tuesday, Wednesday, or Thursday.

At the end of the recruitment period the posters advertising the research were taken down and the registration website taken off-line. Participants were then assigned to their respective groups and informed, by email or SMS, of their acceptance into the study. The initial contact message also informed subjects of the date, time, and venue of their first session, as well as ways of contacting the researcher if they needed further information.

4.3.2 Group and Subject Allocation

The design of the present experiment required subjects to be allocated to either a control group or one of the two experimental groups.

Group allocation was a two step process. Firstly, subjects were allocated to the control group if they had indicated that they were available to participate on a Wednesday afternoon, this was assumed to not be connected with any relevant variable. Specifically, the fact that we were recruiting subjects from any year of study, along with the fact that sessions were scheduled in time-slots where there are typically fewer scheduled classes, was determined to be sufficient protection against any systematic bias that may have been introduced through scheduling issues.

Secondly, those subjects who were available to participate on a Tuesday or Thursday afternoon were randomly assigned to either the Standard Tetris or Modified Tetris experimental groups.

The design of the experiment allowed for more control in allocating subjects to the Standard and Modified Tetris groups than the Control group because, except for a small difference in the training task (described in detail below) the procedure for these two experimental groups was identical. This made it possible to administer their pre-tests, post-tests, and training together in the same sessions.

4.4 Measurement Instrument

4.4.1 Technical specifications

The pre- and post-tests were delivered using WebExp2, an Open Source system developed and implemented by the Department of Informatics at Edinburgh University. The system is designed specifically to enable internet-based administration of psychology experiments that require recording accurate timed response data (Keller et al., 2009), a point discussed further below.

WebExp2 is designed using a client-server architecture (Keller et al., 2009) meaning that the system is comprised of two distinct components. The WebExp2 CLIENT component is responsible for presenting the experiment, recording subject responses, and communicating the timed response data back to the

Server. The Server component, on the other hand, serves as a central repository for all experimental data – stimuli, text for display, etc. – and provides persistent storage for the response data received from instances of the client component.

The WebExp2 client and server components are implemented in the Java programming language which has the distinguishing feature of allowing software written in it to be run on any platform that has an implementation of the Java Virtual Machine. In terms of the present experiment, this allowed the server and client components to be run on two different operating systems, while the distributed nature of the client-server architecture allowed the client and server components to be located in different physical spaces. The server component for the present experiment was hosted on a secure server in Johannesburg running a variant of the GNU/Linux operating system. The WebExp2 client, on the other hand, is designed to run within an Internet browser, such as Firefox or Internet Explorer, on the computer being used to administer the experiment. All instances of the client component for the present experiment thus ran locally on the computers in the laboratory.

4.4.2 Issues with "Web Based Testing"

Several problems have been identified with experiments delivered over the internet. These include issues such as the seriousness with which the subjects complete their tasks, being unable to control distractions in the subject's immediate environment, and confirming the subject's identity (Reips, 2002). Although WebExp2 is a web-based technology and the present experiment was *technically* delivered over the internet – the client and server components communicated via the web – it was not, strictly speaking, an *internet* or *web-based* experiment. This is because all interactions with the subjects were conducted in carefully controlled, laboratory conditions and so most of the problems typically associated with web-based experiments are not of concern.

However, there is one issue with web-based experimentation that carries over to the present study, namely, concern over the accuracy of experiments requiring the collection of response time data. When delivering experiments over the internet one typically has very little control over those factors that could potentially affect response time accuracy, such as the underlying system hardware that determines the responsiveness of the software (graphics card, processor, RAM, etc.), which web browser the subject is using (Internet Explorer, Firefox, Google Chrome, etc.), and which other programs might be sharing system resources with the experimental software (Keller et al., 2009). Again, the fact that the present experiment was conducted under controlled conditions meant that we were able to ensure that all machines running the client software were identical and that there were no other system intensive processes running concurrently. Furthermore, WebExp2 has been tested and shown to yield accurate response data across a number of different conditions. For instance, the experiments presented in Keller et al. (Keller et al., 2009) show WebExp2's timing to be accurate across several different platforms and under different levels of "system load". The study also shows, by replicating a pre-existing psycholinguistic experiment in self-paced reading, that WebExp2 is able to produce results that are comparable to those produced using proprietary experimental hardware and software (Keller et al., 2009). Importantly, using a confidence interval approach, Keller et al. estimate that WebExp2's sensitivity is such that it is able to capture response time data, key-presses in their study, accurate enough to allow the detection of significant differences in RT means as small as 182 ms (Keller et al., 2009).

4.4.3 Customisation of WebExp2: Creating a Mental Rotation Test

The WebExp2 software suite comes with a pre-packaged module for assessing Mental Rotation speed and accuracy. However, this default module was found to be insufficient for the present experiment for two reasons. First, we needed to control the stimuli that were to be presented to the subjects. Secondly, as the module was merely a demonstration of a Mental Rotation paradigm, it only presented a subject taking the test with twenty trials. It was determined that a new WebExp2 module should be developed based on the Shepard/Metzler style Mental Rotation tasks in the Sims and Mayer(2002) study.

It will be recalled that a Shepard/Metzler style MR test presents the comparison and target images, which are either identical or mirror images of one another, side-by-side and never at the same orientation. The subject is then required to determine whether the comparison image is merely a rotated version of the target, or if it has been rotated and also reflected. In the present study, subjects were required to indicate their decision by using their computer's mouse to click one of two buttons marked "Rotated Only" and "Reflected", displayed directly underneath the target and comparison images.

As in the Sims and Meyer (2002) study, four classes of stimuli were used. The

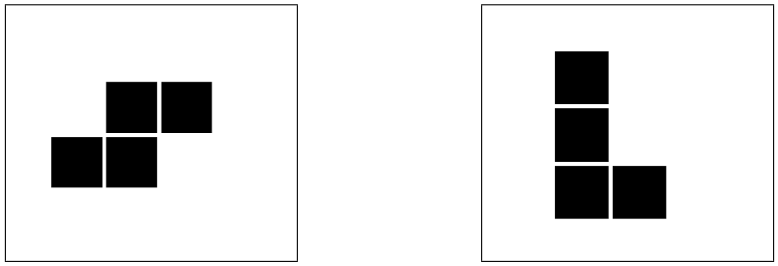

Figure 8: Tetris Shape stimuli

Figure 9: Tetris-Like Letters stimuli

first class of stimuli contained Tetris Shapes, shapes that actually occur within the game itself. Because of the nature of the task, only those Tetris shapes that are not symmetrical were used. The second class of stimuli, the Tetris-Like Letters, was comprised of four shapes, the letters "Z" and "L" and their mirror reflections. These were chosen because of their resemblance to the first class of stimuli. The third class, the Non-Tetris Shapes, bear a strong resemblance to the Tetris shapes although they do not occur in the game. The fourth and final class of stimuli were the Non-Tetris-like Letters, these were the letters "G" and "R" as well as their reflections. These were selected because they do not resemble Tetris shapes. In addition to these four classes of stimuli, a fifth class was created for the demonstration phase of the test. This class's stimuli was comprised of an image of the number "2" and its mirror reflection.

The images used as a basis for the above stimuli classes were initially created using the GNU Image Manipulation Program. These images were then run through a script, bundled with WebExp2, that when provided with a target image produces, firstly, its mirror image and, secondly, a series of images where both the target and its mirror image are rotated around their local origins at increments of 45°. The result of this process was 16 distinct images per, and inclusive of, each target image.

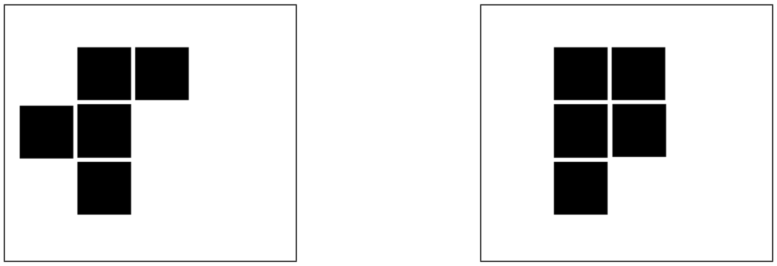

Figure 10: Non-Tetris Shape stimuli

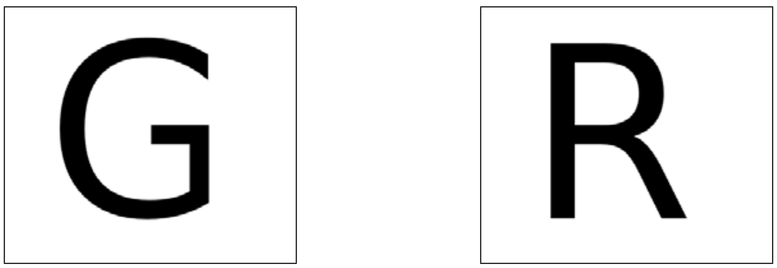

Figure 11: Non-Tetris-like Letter stimuli

Creating a WebExp2 module essentially consists of describing a series of "slides". Each slide is a specification of both what appears, visually or aurally, to the subject during testing as well as the kind of data that is collected by that slide. The present experiment consisted of four sets, or phases, of slides.

The first phase displayed the instructions for the Mental Rotation Test (see Appendix E) as well as collecting the subject's full name and student number. The information collected in this phase was then used to identify all subsequent data collected during the MR test. The second phase consisted of a randomised set of 10 pairs drawn from the demonstration stimuli. While the demonstration phase was timed, all data from this phase was ignored in the data analysis phase. The final screen of the second phase consisted of an informational slide used to inform the subject that their test was about to begin in earnest, that all subsequent trials' responses would be recorded, and that they were urged to respond as quickly and as accurately as they could. The third phase of the test consisted of 160 pairs of images selected from the four classes of stimuli.[19] While

[19]The fact that 160 trials were used was due to two factors. First, we wanted to capture at *least* as much data as Sims and Mayers (2002) who presented subjects with 112 trials. Secondly, due to a feature of WebExp2's image randomization, in order to *guarantee* that we displayed each trial pair at least once, we needed to split the trials into two display groups of 80 trials each.

the ordering of stimuli presentation in this phase was random, the presentation of images was set up in such a way that each pair of images would be displayed at least once, and at most twice. The final phase consisted of a single slide informing the subject that the test was complete and that their responses were being sent back to the server.

4.5 Apparatus and Materials

4.5.1 Computer systems

All pre-tests, post-tests, and training tasks were conducted in the same computer laboratory for all three groups using the same set of computers. The specifications of these laboratory computers were as follows:

- Operating System: Microsoft Windows XP with Service Pack 3.
- Display: 17 inch LCD Monitor running at a resolution of 1360 x 768.
- Intel Core 2 E8300 Processor (2.83 GHz)
- RAM: All machines had 2GB of memory.
- Network: All machines had at least 100Mbit/sec connections to the UKZN Local Area Network.
- Internet Connectivity: Through UKZN's LAN all laboratory computers had access to the internet. The speed of this connection was not guaranteed, but did not impact either pre- and post-tests or training tasks.

The computer that hosted the WebExp2 Server component, as well as all other training task programs, had a GNU/Linux based Operating System running a 3.2GHz Intel Xenon Quad core processor with 4GB of RAM.

4.5.2 Materials used in orientation, pre-tests, and post-tests

Three large cardboard displays were created as visual aids for the orientation, pre-tests, and post-tests. The first was a large board listing all website URLs that were used in the study. This was created in order to avoid confusing subjects by reading out website addresses. At those times when subjects were required to direct their internet browsers to particular URLs the appropriate URL could simply be pointed out on the board rather than having to be spelled out.

The second and third displays were used during orientation to help demonstrate the differences between rotation and reflection. Both boards had a pair of large cardboard cut-outs of the number "2" fixed to them, each taking up roughly one half, left and right respectively, of the board. The board itself was presented in its landscape orientation in order to mimic the dimensions of the computer screen. On both boards the cardboard image on the left was the number "2" at its usual orientation. On the board labelled "Rotated Only" the right side image was a cut-out of the number "2" that was able to be rotated around its center as it was fixed to the board with a drawing pin. This was used to illustrate the instance where the two images were identical except for rotation around one of the image's local origins.

The image on the right hand side of the board labelled "Reflected" was a mirror image of the number "2" and was also able to be rotated. This was used to illustrate instances where the two images were both at different orientations and mirror images of one another.

All boards were given prominent placement and were visible to subjects for the entire duration of the study.

4.5.3 Customised Tetris implementation used in training task

Given the requirements of the experimental design, it would have been impossible to use a pre-existing version of Tetris. Several non-standard elements, described below, were needed to support the present study. In was decided that building a new version of the game from the ground up would have been impractical given constraints on time and resources even though it would have afforded the most control over the end product. The route that was taken was to find an existing version of Tetris that it would be possible to customise to support the non-standard elements. This route had the advantage of significantly bootstrapping the development of the system. The primary disadvantages to this approach was that there was no control over the system's architecture, potentially making factoring the requirements into the game a lot more challenging than if they were built into the system from the beginning.

The version of Tetris chosen for customisation was "JSTetris", originally written by Czarek Tomczak.[20] This version was selected for a number of reasons. Firstly, it was implemented in the JavaScript language and was developed specif-

[20]Although not strictly required (by the licensing agreement under which JSTetris is released) - Tomczak was contacted for permission to use his work in the present study, to which he agreed.

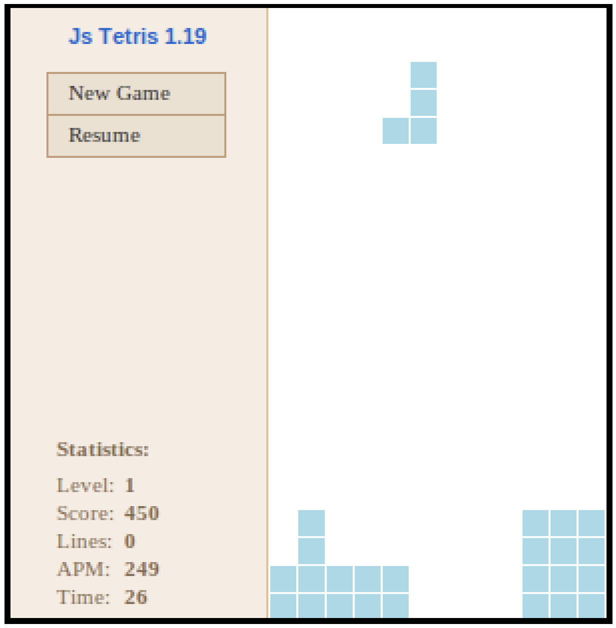

Figure 12: Screen shot of the customised version of JSTetris used for training experimental groups.

ically for deployment over networks using standard internet technologies. This meant that deploying it in the present study would require almost no administrative overhead, such as installing software on every computer in the laboratory. Secondly, the source code for the game is released under the very permissive BSD license, meaning that the software is free to be used and altered under almost any circumstances.[21]

Removal of Visual Clues: It is a standard practice in Tetris implementations to colour or texture zoids based on their type. Given that the present study is focused on the *training of Mental Rotation* this kind of clue presents a potential problem. This is because it is known that as far as object recognition is concerned, if stimuli contain "orientation free" features, such as Hochberg and Gellman's (1977) *landmarks*, the brain is often able to skip Mental Rotation entirely and use these features in determining the stimulus's identity (Takano and Okubo, 2006). Removing this particular visual clue was accomplished by colouring all zoids a single shade of blue. Further, like many Tetris implemen-

[21] see *http://en.wikipedia.org/wiki/BSD_licenses* for an explanation of the license agreement.

tations, JSTetris presents its players with a preview of the next zoid that will enter the game-area – this was removed.

Recording of in-game data: Part of the point of the present study was to investigate epistemic-actions, particularly over-rotation. In order to make this possible it was essential for us to keep records of in-game data detailed enough for us to, ideally, be able to recreate/replay entire games.

In order to achieve this, a series of modifications were made to JSTetris allowing us to capture the following data:

1. For each Game
   (a) The date and time a particular game began. This date was read off the server's internal clock which was synchronised with international date-time servers.
   (b) The student number of the subject playing the game.
   (c) The final score achieved by the subject.
2. For every Episode
   (a) The episode's zoid type.
   (b) The initial orientation of the zoid as it entered the game-area.
   (c) The final orientation and position of the zoid at the end of the episode.
   (d) The final *state* of the game-area at the end of the episode. That is, a snapshot of exactly which parts of the game-area were empty and which parts were filled with zoids or remains of zoids.
3. For every Move within each episode
   (a) The key pressed (rotate, translate left, translate right, drop the zoid)
   (b) The time the key was pressed, measured in milliseconds since the beginning of the episode. Timing of moves was taken care of on the client side (on the computers the subjects were actually using) in order to prevent any timing errors caused by delays in communications back to the server.
   (c) The exact location and orientation of the zoid at the time of the key-press.

At the end of every episode our modified version of Tetris would contact a program on the Server side with the data that had been recorded and this raw data would be written to a database.

Two modes of Rotation: Standard Tetris allows players to rotate their zoids as many times as they want while the zoid is still falling freely. The present study required that one of the groups, the Modified group, be trained on a version of Tetris that only allowed the falling zoid to be rotated until it returned to the orientation at which it entered the game-area. In the case of the T-shapes and L-shapes, this meant that the zoid would be rotated once, 360°, around its local origin, while the Z-shapes and Line-shapes would only be able to be rotated halfway around their local origins. Rotating the Square-shape zoid does not change its orientation.

JSTetris was modified so that the researcher could set whether its game-play mode would be standard Tetris, allowing as many rotations as the subject wanted to make, or in its modified state, where over-rotations were suppressed. When a subject started the game its mode would be set based on the group that subject had been assigned to.

4.5.4 Control Task - Lemmings

The puzzle game "Lemmings" was assigned to the control group as a filler task. The object of the game is to guide a hoard of mindless creatures, the "Lemmings" of the title, through obstacle courses of increasing difficulty. This is accomplished through assigning various roles to individual Lemmings that enable them to alter the landscape of the obstacle courses in order to create a safe path for the rest of the Lemmings.

Lemmings was selected as a control task because it is reasonably engaging and easy to learn. Further, as previous research has shown, it is primarily action-video games that have an effect on perceptual learning (Green and Bavelier, 2006). Lemmings was chosen precisely because it is not an action video-game and, more specifically, because it does not require its players to engage in any tasks that would require Mental Rotation. It was therefore not expected to affect our dependent variable of interest, namely, Mental Rotation performance.

4.6 Procedure

An initial contact SMS and email was sent out to all potential participants two weeks before the first session alerting them that the study was about to begin and that they would start receiving daily reminders about their scheduled sessions. This initial contact SMS was repeated a week before the study began.

Once the experiment had begun, subjects were sent a reminder, by SMS and email, of the date, time, and location of all session they were scheduled to participate in. These reminders were sent both the day before, and the morning of, all sessions, including all pre-tests, post-tests, and training sessions. In almost all communications subjects were told how they could get in contact with the researcher, by email and telephone, if they had any questions about the study or if they could foresee any problems with attending particular sessions.

4.6.1 Orientation and Pre-test administration

The procedure for orientation and pre-test administration was identical across all three sessions (Tuesday, Wednesday, and Thursday) and all three groups (Control, Standard Tetris, and Modified Tetris). All subjects were seated at one of the computers in the laboratory and told that this would be their assigned computer for the duration of the study. Once seated, subjects were provided with consent forms which they were required to read and, if they chose to participate further, sign.

After being given a short introduction to the study as well as the opportunity to ask questions, all groups were read the text introducing the pre-test and explaining what would be required of them (see Appendix D). The pre-test consisted of the Mental Rotation test described in section 4.4.3. In order to access the pre-test, subjects were told to open the web browser Mozilla Firefox and direct it to the now defunct website http://mr.gamesforscience.co.za. Once the page had loaded, the WebExp2 module described above was loaded and the subjects were taken through the four phases described in 4.4.3. Once the pre-test was completed and the subjects' data recorded to the server, subjects' internet browsers were automatically redirected to a page telling them to wait quietly for further instructions from the researcher.

When all subjects in a session had completed their pre-tests, the session then moved into the training phase described directly below. None of the pre-tests took longer than 20 minutes to administer.

4.6.2 Experimental Groups - Standard and Modified Tetris

The Standard and Modified Tetris groups were tested and trained together. Both groups were split across the Tuesday and Thursday sessions. All sessions were held in the same computer laboratory as the pre-tests.

Registers were taken for all sessions. During the first practice session, directly following from the pre-test, subjects were given a briefing describing their training task and what was required from them (see Appendix D). Following this introduction to Tetris, subjects were asked to log into the game delivery system, at which stage they were presented with their assigned version of Tetris, depending on group allocation.

Each practice session consisted of an hour of Tetris play. At the end of each session a lucky draw was held in which a randomly selected subject would win either a R100 gift voucher or an Apple iPod portable MP3 player.

There were five training sessions in total, meaning that each subject in the experimental groups played at least five hours of Tetris.[22] All subjects were asked to refrain from playing any other version of the game for the duration of the study.

4.6.3 Control Group

At the first session, Subjects were given a brief introduction to their task (see Appendix D). Once attendance registers had been taken, control group subjects were required to log into the game delivery system. Here they were asked a series of questions regarding their preferences about money. These data were being used for a pilot study into temporal discounting of monetary amounts. Once these questions were completed, subjects' internet browsers were automatically redirected to their game, Lemmings. They would then play their game for approximately an hour at the end of which a lucky draw was held, as with the

[22]In order to encourage further training on the cognitive task, a competition was run in which a further two iPods were on offer to those subjects who completed an extra five hours of game-play outside of the official training times. Time logs were distributed to all subjects and those interested in participating in this further training were told to record any additional time they spent playing Tetris in order to be eligible for the prize. Unfortunately, this option proved wildly unpopular and none of the subjects completed the task. Further, only two subjects, one from the Standard Tetris group (approx. 1 hour 15 minutes extra) and one from the Modified Tetris group (approximately 47 minutes extra), logged any time over the standard 5 hours of training. From an analysis of these subjects' data it was determined that the impact of what little extra training they did receive was negligable and that it should not affect the results below. We mention this failed attempt at encouraging further participation for the sake of completeness.

experimental groups, in which a randomly selected subject would receive a R100 gift voucher or an Apple iPod.

5 Results

5.1 Sample Characterisation

Table 1 shows that of the 71 individuals who initially signed up to participate in the present study 52 were present for pre-test administration and participated in at least one training session. Only data from subjects who completed all training sessions and both pre- and post-tests were included in the present analysis. This means that a total of 12 subjects who completed the pre-test were excluded, yielding a final sample of 40 subjects and a total attrition rate of 23%. The attrition rates for the Standard (31.5%) and Modified groups (23.5%) were both slightly higher than that of the Control group (14.2%). The median age of the sample was 20 with a standard deviation of 2.65 years. Ages ranged from 19 to 31. A breakdown of median age and age range by group is given in Table 2.

Group	Initial Selection		Pre-test		Post-test	
Control	20	28.2%	16	30.8%	14	35%
Standard Tetris	26	36.6%	19	36.5%	13	32.5%
Modified Tetris	25	35.2%	17	32.7%	13	32.5%
Totals	71		52		40	

Table 1: Group makeup at each stage of the study – presents the number of participants per group as well as the percentage of the total number of participants represented by the group.

Group	Median (Std. Dev)	Range
Control	20 (2.73)	19 – 29
Standard	21 (3.39)	19 – 31
Modified	20 (1.55)	19 – 23
Total	20 (2.65)	19 – 31

Table 2: Summary of subjects' ages by group

Table 3 presents a breakdown of group composition by sex and video-game play experience. Here "Avid Gamers" are those participants who indicated that they generally spent more than 3 hours a week playing video-games, while "Novice Gamers" are those participants who indicated they generally played fewer than 3 hours of video-games per week. There were slightly more male subjects, with females comprising 40% of the total. The male subjects spent more of their leisure time playing video-games than their female counterparts, with males comprising 85% of the avid-gamers group, a distribution consonant with previous research on the relationship between gender and preference in leisure time allocation (Cherney and London, 2006).

Group	Avid Gamers				Novice Gamers				Total Players		
	Sex				Sex				Sex		
	M	F	Totals		M	F	Totals		M	F	Totals
Control	7	0	7	50%	3	4	7	50%	10	4	14
Standard Tetris	5	0	5	38.46%	2	6	8	61.53%	7	6	13
Modified Tetris	5	3	8	61.54%	2	3	5	38.46%	7	6	13
Totals	17	3	20	50%	7	13	20	50%	24	16	40

Table 3: Detailed breakdown of subjects' game play experience by group assignment (Control, Standard, and Modified) and sex (M/F).

Group	Black African				Indian				White				Total Players		
	Sex				Sex				Sex				Sex		
	M	F	Totals		M	F	Totals		M	F	Totals		M	F	Totals
Control	4	1	5	35.71%	2	2	4	28.57%	4	1	5	35.71%	10	4	14
Standard Tetris	5	6	11	84.61%	0	0	0	0%	2	0	2	15.38%	7	6	13
Modified Tetris	6	4	10	76.92%	0	1	1	7.69%	1	1	2	15.38	7	6	13
Totals	15	11	26	65%	2	3	5	12.5	7	2	9	22.5%	24	16	40

Table 4: Group composition by race, group assignment (Control, Standard, and Modified) and sex (M/F)

5.2 Research Question 1 - The effect of training on Mental Rotation ability

Our first research question addressed what effect, if any, the three training programmes corresponding to our group allocation – Control, Standard Tetris, and Modified Tetris groups – would have on subjects' MR performance. We present two analyses of the data, the first was conducted using a one-way Analysis of Covariance with group allocation as the independent variable of interest, the post-test RT data as the dependent variable, and the pre-test RT data as the covariate. The second a was one-way Analysis of Variance on the difference scores between pre- and post-test Response Times.

5.2.1 Data preparation and profile

The data for the present analysis was a strictly defined subset of the total Response Time dataset that was collected during pre- and post-testing. This subset was defined by a set of exclusionary criteria that have become more or less standard practice in the analysis of data for computerised tests of MR ability.

The first of these criteria is only to use data in which the target and comparison images are identical except, of course, in terms of orientation, a practice established by the original Shepard-Metzler (1971) protocol. The reason for this exclusion is because – unlike "rotated" images that are simply rotated around their origin – non-identical, "reflected", images have no straightforwardly specifiable transformation that will bring them into congruence. In consequence, it is not possible to define a simple function that will relate Reaction Time and angular disparity between the two images (Shepard and Metzler, 1971). The second criteria was that all errors, that is misidentification of whether an item was "similar" or "reflected", were excluded. There are a number of reasons for excluding error data but, perhaps, the simplest would be because, in error cases, it's impossible to tell if the subject is actually performing the required task or just answering randomly. Table 6 shows that, on the whole, subjects made very few errors. Finally, all data from the test round were excluded.

Table 5 presents the mean pre- and post-test Response Times for all three groups – as was expected, all three groups show substantial improvements from pre to post-tests.

	Pre-test		Post-test		Difference Scores	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Control	2555.32	654.49	1855.92	372.19	699.37	476.31
Modified	3873.12	1212.99	2487.09	639.96	1386.03	836.94
Standard	3256.62	846.09	2227.71	437.83	1029.65	854.59

Table 5: Mean pre-test, post-test, and difference score RT data per group (ms)

	Pre-test		Post-test	
	Avg. Errors	Std. Deviation	Avg. Errors	Std. Deviation
Control	0.046	0.029	0.035	0.034
Modified	0.045	0.035	0.051	0.042
Standard	0.045	0.034	0.042	0.041

Table 6: Pre- and post-test mean errors per groups

For the one-way ANOVA on difference scores a new variable, CHANGE_RT, was calculated for each subject by subtracting their pre-test Response Time from their post-test Response Time, group means are presented in Table 5.

5.2.2 Testing parametric assumptions

The data were then checked to ensure that there were no violations of the assumptions required by ANCOVA, namely, normality, linearity of the relationship between covariate and independent variable, homogeneity of variances, and homogeneity of regression slopes. ANCOVA requires, in addition to these assumptions, that the covariate's measurement be unaffected by the experimental manipulation and that there be no correlations among covariates. The latter is not an issue in the present analysis because we are only using a single covariate, while the former – that the covariate's measurement be conducted before treatment – forms part of the research design. In this case, independence of the measurement of the covariate holds because our covariate, pre-test RT, was measured before subjects were exposed to their respective experimental manipulations.

Given the relatively small sample size the Shapiro-Wilk test of normality was deemed appropriate for testing whether the Dependent Variable, the post-test RT, was normally distributed (Razali et al., 2011). The results of the Shapiro-Wilk test ($p > 0.5$, see Table 11 in Appendix A), along with visual examination of QQ–plots of the data revealed that the assumption of normality should not be rejected for the Dependent Variable.

Homogeneity of regression slopes can be tested statistically in SPSS by running a preliminary ANCOVA with a custom model that includes covariate by independent variable interaction and checking whether the interaction is significant – if it is not then the assumption holds. The interaction between covariate and Independent Variable was shown not to be significant in our data ($p = .222$), and so we can assume homogeneity of regression slopes.

The assumption of homogeneity of variances was shown not to have been violated using Levene's Test of Equality of Error Variance ($p = .201$).

Assessment of a linear relationship between the covariate, pre-test RT, and the Independent Variable, the post-test RT, was accomplished through the visual examination of a scatter plot (see Figure 13 in Appendix A). Determining that there is a *definite* linear relationship between the pre- and post-tests for every group is challenging with so few data-points. However, none of the three groups demonstrated any clear evidence for *non-linearity*, and so our assumption of linearity is satisfied.

The one-way ANOVA on difference scores assumes that the dependent variable of interest be approximately normally distributed for each category of the independent variable. Again we tested this assumption by running the Shapiro-Wilk test on each of the group's difference scores, revealing that the assumption of normality need not be rejected ($p > 0.5$, see Table 12 in Appendix A). The one-way ANOVA further requires that the assumption of homogeneity of variances holds between the independent groups. This assumption was not rejected as Levene's Test of Equality of Error Variance ($p = .093$) failed to reach significance.

5.2.3 Research Question 1 core results

With preliminary checks completed the one-way between-groups ANCOVA was conducted, results are shown in Table 13 in Appendix A. After adjusting for pre-test Response Times there was no significant difference found between the three intervention groups on post-test RT, $F(2,36) = .695$, $p = 0.506$. The Covariate, pre-test RT, was significantly related to the participants' post-test RTs, $F(1,36) = 21.5$, $p < .001$.

The one-way ANOVA on difference scores' results are shown in Table 14, Appendix A. No statistically significant differences were found between groups on their difference scores between pre- and post-test Response Times ($p = .066$).

5.3 Research Question 2 - Does over-rotation of Tetris zoids increase with training?

Kirsh and Maglio (1994) argued that if all Tetris players' in-game actions were purely pragmatic – that is, if every action was undertaken *solely* to move them closer to some end state – then we should expect the average number of 90°rotations to be roughly half the number of rotation operations required to get the zoid back into the orientation it was in when entering the playing field – the "expected" number of rotations per zoid type are shown in the final column of Table 7. For instance, the L-shaped zoid required four 90° rotations to bring it back to its original orientation, so we should expect the average number of 90° rotations to be 1.5. However, the game play data that they had collected showed their Tetris players rotating more often than this and it was these extra rotations that provided part of their evidence for the existence and use of epistemic actions by Tetris players. Although Kirsh and Maglio fail to provide the exact numbers – much of their data is presented only in graphical form – Table 7 also presents an approximation of the average number of rotations per zoid type as presented in Figure 6 of (Kirsh and Maglio, 1994).

Zoid Type	Kirsh and Maglio	Standard Tetris	Expected Avg.
L-Shape	1.80	1.86	1.50
T-Shape	1.70	1.68	1.50
Z-Shape	0.70	0.70	0.50
Line	0.58	0.54	0.50
Square	0.02	0.01	0.00

Table 7: Approximate average rotations per zoid type in Kirsh and Maglio (1994) compared with average rotations per zoid type for the Standard tetris group.

Following Kirsh and Maglio's demonstration of over-rotation, we examined our Standard Tetris group's in-game data in a similar manner. Table 7 shows the actual average number of 90° rotations during our subjects' training period. It is clear that our Standard Tetris group shows the same pattern of over-rotation as Kirsh and Maglio's subjects, that is, according to their criteria our subjects seem to be over-rotating their zoids.

The second aim of our study was to investigate whether the Standard experimental group's average number of rotations *changed* significantly over the course of their training. This analysis was accomplished through the use of a series of paired sample t-tests comparing the average number of rotations per zoid type for the Standard Tetris group's first ten games with the average number of rotations per zoid type for their last ten games.

5.3.1 Data preparation and testing Parametric assumptions

As has been mentioned, the data for the present analysis were drawn from the first and last ten games of Tetris played by subjects allocated to the Standard Tetris experimental group. For every episode the total number of successful rotations were calculated. These were then used to calculate each subjects' average number of rotations *per zoid type* at the beginning (first 10 games) and end (last 10 games) of their training period, yielding the dataset shown in table 8.

Taking our data from the first and last 10 games was justified by the fact that, although the total number of games our Standard Tetris subjects played varied considerably (average = 59.77, SD = 10.88), each of them played at least 10 games in both their first and last training sessions.[23] Limiting our analysis

[23]The Modified Tetris group, on the other hand, played an average of 61.92 games (SD = 29.61). It is important to note that subjects in both groups were, on average, exposed to

to these 20 games at the extremes of the training period ensured that we only analysed data drawn from the first and last training sessions.

	Line-Shape		T-Shape		Z-Shape		L-Shape		Square-Shape	
Subj.No	First	Last	First	Last	First	Last	First	Last	First	Last
21	0.6181	0.5536	1.4898	1.5404	0.6578	0.5691	1.9323	1.9171	0.0078	0.0162
22	0.5124	0.5319	1.9912	2.7108	0.7679	0.5414	2.2333	2.4677	0.0081	0.0000
24	0.4792	0.6308	1.12	1.5680	0.5810	0.7793	1.5359	1.9245	0.0085	0.0072
28	0.5385	0.5625	1.8583	1.2222	0.6107	0.3731	1.8667	1.4848	0.0000	0.0000
29	0.4712	0.5462	1.9846	2.0398	0.7506	0.7224	1.9052	1.9594	0.0000	0.0061
31	0.4714	0.4649	1.8736	1.3760	1.5221	0.8807	2.0313	1.6471	0.0556	0.0101
34	0.6429	0.5294	1.6282	1.6963	0.6708	0.6216	1.7114	1.8460	0.0460	0.0466
39	0.5395	0.6529	1.7023	2.0516	0.9280	0.7394	1.8373	2.0393	0.0395	0.0091
43	0.4091	0.4405	1.2459	1.3452	0.7607	0.6627	1.1008	1.5385	0.0364	0.0682
46	0.3134	0.4235	0.5217	0.9552	0.5635	0.6310	1.0432	1.1086	0.0179	0.0125
47	0.5729	0.4444	1.3239	1.2813	0.7891	0.7027	2.0671	1.7059	0.0250	0.0000
90	0.5574	0.5101	1.4400	1.3964	0.5423	0.6424	1.2252	1.7332	0.0152	0.0000
92	0.5189	0.5362	1.7636	1.5068	0.8889	0.8480	2.2000	2.0448	0.0108	0.0000

Table 8: Average number of rotations per zoid type for the Standard tetris group's first and last 10 tetris games.

Although we do not include the Square-shape zoid data in the following analysis – there is not nearly enough data for any serious analysis – it is interesting to note that at least *some* of the time subjects attempted to rotate these zoids even though it has no practical effect in the game.

Paired sample t-tests require that three assumptions about the data hold, firstly, that the data is normally distributed, secondly, that the differences between the two scores obtained for each subject be normally distributed, and, finally, that variances are equal.

Table 16 in Appendix A shows the results of our tests assessing the assumption that our samples, and differences between scores, are normally distributed. Note that the assumption of normality is violated for the variable FIRST_10_Z - that is, it is violated for the beginning average rotations for the Z-shaped zoid data. Examining the data in Table 8 reveals that subject 31's average number of rotations for Z-shape zoids is almost three standard deviations greater than the mean.[24] It may be possible to motivate excluding this data-point from our

more than the estimated 38 games (footnote 11 above) of Tetris required to approximate the number of MR trials within which individuals have been shown to reach asymptotic levels of MR performance (Kail and Park, 1990).

[24]More importantly, given that our analysis is conducted using paired samples t-tests, the

analysis for a number of reasons. We might, for instance, use the fact that this subject does *not* seem to be an outlier with regards to average rotation on other zoid types to argue that this data-point should be excluded. However, given that we are interested in over-rotation it is important for us not to *simply* exclude any case of over-rotation from our dataset. We therefore present a number of different analyses for the Z-shape zoid's beginning and ending average rotation data. Firstly, for completeness, we present but do not discuss a paired sample t-test on the non-normal dataset including subject 31's beginning score. Secondly, we created a new variable FIRST_10_Z_SANS_OUTLIER that, as the name suggests, excludes subject 31's data (see Table 10). This was then used on a separate paired sample t-test comparing beginning and ending Z-shape average rotations (see Table 10). Finally, we ran a non-parametric alternative to the paired samples t-test, namely the Wilcoxon Signed-rank test (see Note 25). The assumption of equality of variances was tested using a series of Levene's tests, none of which reached significance (all $p > .05$).

Subj. No.	Beginning Average Score	End Average Score	Difference Score
21	243712	492220	248508
22	149146	97674	-51472
24	131545	187609	56064
28	400928	92874	-308054
29	541200	1534793	993593
31	45520	147169	101649
34	143429	1044740	901311
39	349717	1319748	970031
43	24931	132214	107283
46	9584	28178	18594
47	119000	63950	-55050
90	64200	986325	992125
92	216534	254350	37816

Table 9: Average beginning and ending tetris scores

5.3.2 Research Question 2 core results

With parametric assumptions in place, a series of paired sample t-tests were run in order to compare the Standard Tetris group's average number of rotations at

difference between subject 31's beginning (1.5221) and ending (0.8807) average rotations for Z-shapes is *more* than three standard deviations greater than the mean paired difference between the rest of the dataset (see Table 10).

the beginning and end of their training period, the results of which are shown in Table 10. [25] From these results we see that there were no statistically significant changes in the average number of rotations for any zoid type from the beginning to the end of training (all $p > .05$).

	Paired Differences							
				95% Confidence interval				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
FIRST_10_LINE - LAST_10_LINE	-.0140000	.0861766	.0239011	-.0660760	.0380760	-.586	12	.569
FIRST_10_T - LAST_10_T	-.0574538	.3794234	.1052331	-.2867371	.1718294	-.546	12	.595
FIRST_10_Z - LAST_10_Z	.1015077	.2050396	.0568678	-.0223965	.2254119	1.785	12	.100
FIRST_10_L - LAST_10_L	-.0559385	.3067704	.0850828	-.2413180	.1294410	-.657	12	.523
FIRST_10_Z_SANS_OUTLIER - LAST_10_Z	.0565167	.1309826	.0378114	-.0267057	.1397390	1.495	11	.163

Table 10: Paired sample t-tests comparing beginning of training average number of rotations with end of training average number of rotations per zoid type.

In order to test whether our subjects' Tetris ability improved over the training period a Wilcoxon Signed-rank test was used to analyse the average scores for their first and last ten games. This revealed a statistically significant increase in Tetris scores (Z = -2.062, p = .039) from the beginning to the end of training, with a large effect size (r = .57).

5.4 Research Question 3

The final aim of our study was to explore any relationships between in-game rotation of zoids, Tetris score, and pre- and post-tests of MR performance. Bivariate correlation was performed to check for correlations between variables, the results of which are shown in Table 15 in Appendix A.

Pre-test MR score was negatively correlated with the average number of rotations of T-shape zoids (r = -.556, $p < .05$) while MR post-test scores were negatively correlated with the average number of rotations of Z-shape zoids, excluding subject 31's results (r = -.686, $p < .05$).

With regards to Tetris performance significant positive correlations were found for beginning Tetris score and average rotation on Line-shape (r = .596,

[25] A Wilcoxon Signed-rank test was run in order to analyse the Z-shape zoid data that was found to violate the assumption of normality. This revealed that there was no statistically significant change (Z = -1.642, p = .101) in the average number of rotations from the beginning to the end of the training period.

$p < .05$) and T-shape ($r = .588$, $p < .05$) zoids.

6 Discussion

6.0.1 Research Question 1

To reiterate, the primary aim of the present research was to investigate whether epistemic actions available to Tetris players have a measurable effect on the game's efficacy as a Mental Rotation training task. In addressing this question we administered our subjects Shepard-Metzler style tests of MR performance both before and after their training periods.

The first point to note about our results is that all three groups experienced improvements in their MR performance from pre- to post-test (Table 5). This result was not unexpected, as it is well established that subjects tend to demonstrate large improvements in Response Times from simply retaking an MR test (Peters et al., 1995).

What is important to note is the fact that the pre-test average RT for the groups differ quite substantially (a limitation on the study which we will discuss in more detail below). Specifically, our control group's pre-test score was approximately 700 ms faster than the Standard Tetris group's score and approximately 1300 ms faster than the Modified Tetris group's average RT score (Table 5). These pre-test differences provide critical information for interpreting pre to post-test difference scores. Although our one-way ANOVA on difference scores approached, but did not reach, statistical significance (p=.066), simple examination of the magnitude of the three groups' difference scores might suggest that our Modified version of Tetris was more effective than Standard Tetris (which was, in turn, more effective than the Control group's task) for training MR, as the Modified Tetris group's gains (1386 ms) were larger than those of the Standard Tetris group's (1029 ms) and almost twice as large as those of the Control group (699 ms).

However, as Kail's work on MR has demonstrated (Kail and Park, 1990), overall improvement in MR performance with practice tends to be *relative* to initial MR ability. Kail has shown that both hyperbolic and power functions are relatively good fits for data recording the improvement on MR performance with practice (Kail, 1986). Both of these functions are characterised by an initial, rather sharp, drop in MR Response Times followed by a gradual leveling out as subjects approach asymptotic levels of performance. Given this pattern of improvement, a group's overall gains in MR Response Time with practice should depend on how close they are to their asymptotic levels of performance at the time of their pre-test. If the differences in our groups' pre-test scores do *in fact* reflect that they were at different distances from their asymptotic levels of performance then, even *without* our particular experimental interventions, it is possible that we would have observed the same pattern in difference scores.

Our primary analysis of the pre-test/post-test RT data was conducted using a one-way ANCOVA with post-test RT data as the dependent variable of interest and pre-test RT data as the covariate. As with our one way ANOVA on difference scores, our ANCOVA revealed no significant differences (p>.05) between our three groups' post-test RT once pre-test RT had been statistically controlled for. The combined results from both our analyses suggest that – beyond the well established MR gain that comes from retesting – there were no systematic improvements to MR performance due to Tetris training in general, and no benefits from playing our Modified version of Tetris in particular. Given that *both* out Tetris groups failed to show greater improvements in MR performance than the non-Tetris control group, we are in the position to draw two tentative conclusions.

Our first addresses the question we posed at the outset of this study – it seems as though the existence of epistemic actions *does not* affect Tetris' efficacy as a Mental Rotation training task – at least not with 5 hours of training. This conclusion is, however, a trivial consequence of our second, which is that – at least insofar as our subjects are concerned – Tetris training seems to not have any benefit as a MR training task *in general*. Our results are not unique either, as they mirror those of Sims and Mayers' (2002) longitudinal study in which they too found no MR improvements with 12 hours of Tetris training.

As we suggested in the discussion of Sims and Mayers' results above, it may still be possible that differential performance improvements might emerge with a *longer* training period. Indeed, we may even find that with a longer training period the existence of epistemic actions *does* have an effect on the efficacy of Tetris as a training task. This possibility raises an important issue about using video-games to improve perceptual and cognitive processes generally. At least part of the motivation for investigating the effects of video-games is to identify tools with which to improve cognition and perception in cases where individuals may be under performing – for example, addressing possible sex-dependent differences in MR performance (Cherney, 2008) – or cases in which certain individuals might benefit from above average performance – for example, Air Force pilots whose work has extremely high cognitive and perceptual demands (Gopher et al., 1994). As we have seen, though, the results of work investigating the impact of Tetris on Mental Rotation has been, if not contradictory then at least inconsistent, with studies demonstrating differential MR improvements in male subjects only (Okagaki and Frensch, 1994), primarily in female subjects (Cherney, 2008), and in no subjects at all (Sims and Mayer, 2002). One way of viewing the present study is to see it as an attempt to design a version of Tetris that had the best chance of eliciting improvements in MR performance, given what we know about the dynamics of the game. Even with these modifications we failed to see any significant improvement in our Tetris playing subjects' MR performance when they were compared to a control group. Although it does seem that under *some* circumstances Tetris might improve MR performance, it is important to consider whether having subjects play several hours of the game – even a version designed *specifically for training MR* – for a *potential* improvement should be considered a *useful* intervention, given that large MR performance gains can be achieved in a short time by simply practicing the MR task itself (Boot et al., 2008).

6.0.2 Research Question 2

We also addressed the question of whether our Standard Tetris group would demonstrate a significant increase in the average number of rotations they used as their Tetris skills improved.

On average, the group's Tetris skill – in terms of number of points per game – improved significantly from the beginning to the end of their training period (p = .039). Examining individual scores we found that all except three of the subjects in the Standard Tetris group showed gains in their average scores (see Table 9). It is not clear why subject numbers 22, 28, and 47's showed losses in their overall performance but an examination of their game score data failed to reveal any extremely high scoring games in their first 10 games that might have pushed their beginning score averages upwards. Further, examination of their scores show that subjects 28 and 47 had a number of extremely low scoring games during their final 10 games. Perhaps the simplest explanation of these results – in absence of any evidence other than the subjects' Tetris scores – is that at the end of the training period these particular subjects were no longer motivated to participate wholeheartedly in the study or were simply bored of the game. We discuss the former possibility further below.

Although our subjects showed improved Tetris skill across the training period, a series of paired sample t-tests comparing the average number of rotations per zoid-type at the beginning and end of the training period failed to yield any statistically significant differences ($p > .05$ for all zoid types).

Recall that we were initially interested in this question because a clear increase in average number of rotations as our subjects' Tetris skills improved would provide us with a *prima facie* reason to rule out any over-rotations we do see in the data as being the result of simple errors and help to justify our interpretation of at least some of those over-rotation *as* epistemic actions. Unfortunately, the fact that we see no clear increase *or* decrease in average rotation means that we are in a slightly weaker position when it comes to interpreting the patterns of over-rotation than if there were such a difference. However, if we still follow Kirsh and Maglio in interpreting at least *some* of the over-rotations that we see in our Tetris players' data (Table 7) as epistemic actions, then it seems as though our subjects demonstrated no increased use of epistemic actions from the beginning to the end of their training period.

There are two important challenges to this conclusion though. Firstly, as we have mentioned, Destefano et al. point out (Destefano et al., 2011) that actually interpreting an over-rotation or translation routine *as* an epistemic action is not straightforward. Given in-game data of the kind recorded in the present study it is not always possible to distinguish between, for example, a subject over-rotating their zoid to match the contour at the bottom of the playing area or a subject who has accidentally pressed the rotate button and now has to over-rotate to compensate for the error. If we assume that beginning Tetris players are more likely to over-rotate their zoids to compensate for errors, while experienced players over-rotate their zoids to reduce mental effort, the fact that there is no significant change in average number of rotations may simply reflect that fact that both beginners and more experienced players over-rotate roughly the same amount, but for different reasons.

Secondly given the small size of our Standard Tetris group (an issue we discuss below) as well as the fact that that Maglio and Kirsh report only very small increases in incidence of epistemic actions with training (Maglio and Kirsh, 1996), we should not rule out the possibility of a Type II error.

6.0.3 Research Question 3

Finally, we were interested in whether there were any correlations between average number of zoid rotations, performance in Tetris, and performance in tests of MR ability. As this question was primarily exploratory there were, as mentioned, no comprehensive expectations from these data, especially in light of the conflicting evidence about the incidence of over-rotation with skill from Maglio and Kirsh (1996) and Destefano et al. (2011).

When approaching these data it is important to look at the pattern of correlations as a whole. Although there were only a handful of significant correlations (these fall in line with the following discussion), what is possibly more important is that there seems to be a pattern of weak correlations between Tetris performance and average number of zoid rotations as well as between Tetris performance and MR performance.

It's interesting to note that the subjects' ending Tetris scores show, on the whole, positive (but weak) correlations with average number of zoid rotations. This is more or less the direction of covariance we would expect, given Kirsh and Maglio's assertion that *expert* Tetris players make use of epistemic actions (Kirsh and Maglio, 1994) and that the use of epistemic actions supposedly increase with players' skill (Maglio and Kirsh, 1996).

What was unexpected was the direction of the correlations we observed between MR performance and rotations. Given that one of the proposed epistemic functions assigned to zoid rotation was simplifying the Mental Rotation tasks in Tetris, we would not have been surprised to see that subjects who performed poorly on their MR tests rely on this class of epistemic action. The pattern of correlations in our data suggest the opposite may be the case, as we observe that our subjects' MR performance scores for *both* pre- and post-tests tend towards being weakly negatively correlated with average number of rotations. Note that a negative correlation between MR performance and average number of rotations for a zoid means that subjects who demonstrate *better* MR performance tend to rotate their zoids *more*. This raises some interesting questions that neither Kirsh and Maglio nor Destefano and his colleagues seem to have considered such as – as seems to be the case with the present study – are subjects who demonstrate better visual-spatial performance in general more likely to use epistemic actions? Or, is there a relationship between use of epistemic actions and intelligence or personality profile? Again, given the size of our Standard group these results are provisional, but interesting, and the associated commentary is speculative at best. Further research with appropriately sized samples is required.

6.1 Issues with the present study and improving subsequent stages of research

Although the recruitment phase proceeded without incident, the overall response to the study was, in general, quite poor. Informal discussions with potential subjects suggest the fact that the study required several weeks of commitment and the fact that the days and time set for participation were limited were the primary reasons they declined the opportunity to participate. Poor response to the recruitment phase along with attrition of the group over the course of the study lead to small final sample sizes, which lead to potential issues with statistical power as well as introducing the potential for outliers to significantly affect our results. Subsequent iterations of the research can address the problems with small sample sizes by simply signing up a larger group of participants. Importantly, the system developed for the present study is fully internet ready and, although there would need to be a much stricter set of criteria for participation in order to deal with the potential problems associated with web based testing (Reips, 2002), it may be possible to recruit a fairly large number of subjects from across the world by advertising the research on social networks.

Delivering the experiments over the internet would also allow us to address the issue of random assignment to groups. Our fixed schedule for testing and training reduced flexibility in group allocation and opened the study up to sampling bias. Particularly problematic was having to assign all subjects available on a Wednesday to the Control group. As we have seen above, this group's pre-test MR Response Times were faster than the two Experimental groups. It is certainly possible that the Control group's pre-test performance is a reflection of an underlying bias in the sample. Fortunately, the fact that the Standard and Modified Tetris groups shared sessions meant that we were able to use randomised assignment when allocating subjects to one or the other. Although this doesn't guarantee that these groups were representative of the student population as a whole, it does help guard against systematic non-equivalence *between* these two groups. This seems to be reflected in data as the pre-test MR Response Times for the two Tetris groups are much closer to one another than the Control group.

A further potential issue with the present study is illustrated by the case of the three Standard group subjects who showed an overall *loss* in their total score between the beginning and end of the training sessions. This raised the possibility that some subjects *may* have experienced an overall decline in motivation to play the game which may have been exacerbated by the fact that subjects were not guaranteed to receive compensation for their participation in the study but were, rather, afforded the possibility of winning a prize every week.

Subsequent longitudinal studies may better motivate subjects to improve their performance by tying their compensation to their achievements in the game. One possibility would be to link subjects' compensation to their scores and, at the end of their training sessions, inform them of how much money they have earned while they played and encourage them to improve in order to increase their rewards in future sessions.

6.2 Conclusion

The aim of the present study was to investigate whether epistemic actions in Tetris impact its effectiveness as a Mental Rotation training task. A group of subjects who received at least 5 hours of Tetris training on a version of the game that made epistemic actions involving rotation impossible showed no greater MR performance gains when their results were compared to a group of subjects trained using a Standard version of Tetris. This suggests that the occurence of epistemic actions *does not* have an impact on Tetris' efficacy as a MR training task. Further, neither of the groups assigned Tetris training showed greater MR performance improvements than a non-Tetris control group, a result that is not unprecedented (Sims and Mayer, 2002) and which suggests that, at least under some circumstances, Tetris training fails to impart MR performance gains any greater than what can be expected due to retest effects.

References

Amponsah, B. and Krekling, S. (1997). Sex Differences in Visual-Spatial Performance among Ghanaian and Norwegian Adults, *Journal of Cross-Cultural Psychology* **28**(1): 81–92.

Bartlett, J. C. and Searcy, J. (1993). Inversion and configuration of faces, *Cognitive psychology* **25**(3): 281–316.

Bethell-Fox, C. E. and Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity, *Journal of Experimental Psychology: Human Perception and Performance* **14**(1): 12–23.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M. and Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control, *Acta psychologica* **129**(3): 387–98.

Casey, M. and Nuttall, R. (2001). Spatial-Mechanical Reasoning Skills versus Mathematics Self-Confidence as Mediators of Gender Differences on Mathematics Subtests Using Cross-National Gender-Based Items, *Journal for Research in Mathematics* **32**(1): 28–57.

Ceci, S. J., Williams, W. M. and Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations., *Psychological bulletin* **135**(2): 218–61.

Cherney, I. D. (2008). Mom, Let Me Play More Computer Games: They Improve My Mental Rotation Skills, *Sex Roles* **59**(11-12): 776–786.

Cherney, I. D. and London, K. (2006). Gender-linked Differences in the Toys, Television Shows, Computer Games, and Outdoor Activities of 5- to 13-year-old Children, *Sex Roles* **54**(9-10): 717–726.

Cooper, L. A. (1975). Mental Rotation of Random Two-Dimensional Shapes, *Cognitive Psychology* **7**: 20–43.

Cooper, L. A. (1976). Demonstration of a mental analog of an external rotation, *Perception & Psychophysics* **19**(4): 296–302.

Cooper, L. A. and Podgorny, P. (1976). Mental transformations and visual comparison processes: effects of complexity and similarity, *Journal of experimental psychology: Human perception and performance* **2**(4): 503–14.

Cooper, L. A. and Shepard, R. N. (1973). Chronometric studies of the rotation of mental images, *in* W. G. Chase (ed.), *Psychological Science*, Vol. 171, Academic Press, chapter 3, pp. 75–176.

De Lisi, R. and Wolford, J. L. (2002). Improving Children's Mental Rotation Accuracy With Computer Game Playing, *The Journal of Genetic Psychology* **163**(3): 272–282.

Destefano, M., Lindstedt, J. K. and Gray, W. D. (2011). Use of Complementary Actions Decreases with Expertise Substituting Actions in-the-world for Processes, *in* L. Carlson, C. Hoelscher and T. Shipley (eds), *Proceedings of the 33rd annual conference of the cognitive science society*, Cognitive Science Society, Austin, TX, pp. 2709–2714.

Dimitrov, D. M. and Rumrill, P. D. (2003). Pretest-posttest designs and measurement of change, *Work (Reading, Mass.)* **20**(2): 159–65.

Donovan, J. J. and Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't, *Journal of Applied Psychology* **84**(5): 795–805.

Dorval, M. and Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization, *Perceptual and motor skills* **62**: 159–162.

Dror, I. E., Schmitz-Williams, I. C. and Smith, W. (2005). Older adults use mental representations that reduce cognitive load: mental rotation utilizes holistic representations and processing, *Experimental aging research* **31**(4): 409–20.

Folk, M. D. and Luce, R. D. (1987). Effects of stimulus complexity on mental rotation rate of polygons, *Journal of experimental psychology: Human perception and performance* **13**(3): 395–404.

Geary, D. C. and Desoto, M. C. (2001). Sex Differences in Spatial Abilities Among Adults from the United States and China, *Evolution and Cognition* **7**(2): 172–177.

Geiser, C., Lehmann, W. and Eid, M. (2006). Separating "Rotators" From "Nonrotators" in the Mental Rotations Test : A Multigroup Latent Class Analysis, *Multivariate Behavioral Research* **41**(3): 261–293.

Goldstein, E. B. (2008). *Cognitive Psychology: Connecting Mind, Research, and Everyday Experience*, 2 edn, Thomson Wadsworth, Belmont.

Gopher, D., Well, M. and Bareket, T. (1994). Transfer of Skill from a Computer Game Trainer to Flight, *Human Factors: The Journal of the Human Factors and Ergonomics Society* **36**(3): 387–405.

Green, C. and Bavelier, D. (2006). The cognitive neuroscience of video games, *in* P. Messaris and L. Humphreys (eds), *Digital media: Transformations in human Communication*, Peter Lang, New York.

Haier, R. J., Karama, S., Leyba, L. and Jung, R. E. (2009). MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task, *BMC research notes* **2**: 174.

Heil, M., Rösler, F., Link, M. and Bajric, J. (1998). What is improved if a mental rotation task is repeated–the efficiency of memory access, or the speed of a transformation routine?, *Psychological research* **61**(2): 99–106.

Hochberg, J. and Gellman, L. (1977). The effect of landmark features on mental rotation times, *Memory & cognition* **5**(1): 23–6.

Hooven, C. K., Chabris, C. F., Ellison, P. T. and Kosslyn, S. M. (2004). The relationship of male testosterone to components of mental rotation, *Neuropsychologia* **42**(6): 782–90.

Jansen-Osmann, P. and Heil, M. (2007). Suitable stimuli to obtain (no) gender differences in the speed of cognitive processes involved in mental rotation, *Brain and cognition* **64**(3): 217–27.

Just, M. A. and Carpenter, P. (1976). Eye fixations and cognitive processes, *Cognitive Psychology* **8**(4): 441–480.

Kail, R. (1986). The impact of extended practice on rate of mental rotation., *Journal of experimental child psychology* **42**(3): 378–91.

Kail, R. and Park, Y. S. (1990). Impact of practice on speed of mental rotation, *Journal of experimental child psychology* **49**(2): 227–44.

Keller, F., Gunasekharan, S., Mayo, N. and Corley, M. (2009). Timing accuracy of Web experiments: a case study using the WebExp software package, *Behavior research methods* **41**(1): 1–12.

Kirsh, D. and Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Action, *Cognitive Science* **18**(4): 513–549.

Levine, S. C., Huttenlocher, J., Taylor, A. and Langrock, A. (1999). Early sex differences in spatial skill, *Developmental psychology* **35**(4): 940–9.

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S. and Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill, *Psychological science* **16**(11): 841–5.

Linn, M. C. and Petersen, A. C. (1985). Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis, *Child Development* **56**(6): 1479.

Maglio, P. (1995). *The Computational Basis of Interactive Skill*, PhD thesis, The University of California, San Diego.

Maglio, P. and Kirsh, D. (1996). Epistemic Action Increases With Skill, *Proceeding sof the 18th Annual Meeting of the Cognitive Science Society*, Erbaum, Mahwah, NJ, pp. 391–396.

Maglio, P. P. and Wenger, M. J. (2000). Two Views are Better than One : Epistemic Actions May Prime, *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Mahwah, NJ.

Maglio, P. P., Wenger, M. J. and Copeland, A. M. (2008). Evidence for the role of self-priming in epistemic action: expertise and the effective use of memory, *Acta psychologica* **127**(1): 72–88.

Malone, T. W. (1983). How Do People Organize Their Desks? Implications for the Design of Office Information Systems, *ACM Transactions on Office Information Systems* **1**(1): 99–112.

Mann, V., Sasanuma, S. and Sakuma, N. (1990). Sex differences in cognitive abilities: A cross-cultural perspective, *Neuropsychologia* **28**(10).

Masters, M. S. (1998). The gender difference on the Mental Rotations test is not due to performance factors, *Memory & cognition* **26**(3): 444–8.

McClurg, P. and Chaille, C. (1987). Computer Games: Environments for developing spatial cognition?, *Journal of Educational Computing Research* **3**: 95–111.

Moore, D. S. and Johnson, S. P. (2008). Mental rotation in human infants: a sex difference, *Psychological science* **19**(11): 1063–6.

Moore, D. S. and Johnson, S. P. (2011). Mental Rotation of Dynamic, Three-Dimensional Stimuli by 3-Month-Old Infants, *Infancy* **16**(4): 435–445.

Okagaki, L. and Frensch, P. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence, *Journal of Applied Developmental Psychology* **15**(1): 33–58.

Ozel, S., Larue, J. and Molinaro, C. (2004). Relation Between Sport and Spatial Imagery: Comparison of Three Groups of Participants, *The Journal of Psychology* **138**(1): 49–63.

Pallant, J. (2011). *SPSS Survival Manual*, 4th edn, Allen & Unwin, New South Wales: Australia.

Peters, M. and Battista, C. (2008). Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library, *Brain and cognition* **66**(3): 260–4.

Peters, M., Laeng, B., Latham, K. and Jackson, M. (1995). A redrawn Vandenberg and Kuse mental rotations test - different versions and factors that affect performance, *Brain and cognition* **28**: 39–58.

Pinker, S. (1998). *How the mind works*, Penguin, London.

Pylyshyn, Z. (1979). The rate of "mental rotation" of images : A test of a holistic analogue hypothesis, *Memory & Cognition* **7**(1): 19–28.

Pylyshyn, Z. W. (2002). Mental imagery: in search of a theory, *Behavioral and Brain Sciences* **25**(2): 157–82; discussion 182–237.

Quinn, P. C. and Liben, L. S. (2008). A sex difference in mental rotation in young infants, *Psychological science* **19**(11): 1067–70.

Razali, N. M., Wah, Y. B. and Sciences, M. (2011). Power comparisons of Shapiro-Wilk , Kolmogorov-Smirnov , Lilliefors and Anderson-Darling tests, *Journal of Statistical Modeling and Analytics* **2**(1): 21–33.

Reips, U. (2002). Standards for Internet-Based Experimenting, *Experimental Psychology* **49**(4): 243–256.

Riesberg, D. (1987). External Representations and the Advantages of Externalizing One's Thoughts, *Proceedings of the 9th Annual Conference of the Cognitive Science Society*, Morgan Kaufman.

Shepard, R. and Metzler, J. (1971). Mental rotation of three-dimensional objects, *Science* **171**(3972): 701–703.

Shepard, R. N. (1978). The Mental Image, *American Psychologist* (February).

Sims, V. K. and Mayer, R. E. (2002). Domain specificity of spatial expertise: the case of video game players, *Applied Cognitive Psychology* **16**(1): 97–115.

Smith, W. and Dror, I. E. (2001). The role of meaning and familiarity in mental transformations, *Psychonomic bulletin & review* **8**(4): 732–41.

Sternberg, R. and Sternberg, K. (2011). *Cognitive Psychology*, 6th edn, Wadsworth Publishing, London.

Takano, Y. and Okubo, M. (2006). Mental Rotation, *Encyclopedia of Cognitive Science*, John Wiley & Sons, Ltd.

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

Thompson, P. (1980). Margaret Thatcher: a new illusion, *Perception* **9**: 483–484.

Voyer, D. and Sullivan, A. (2003). The relation between spatial and mathematical abilities: Potential factors underlying suppression, *International Journal of Psychology* **38**(1): 11–23.

Voyer, D., Voyer, S. and Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables, *Psychological bulletin* **117**(2): 250–70.

Wallop, B. (2009). Video games bigger than film.
URL: *http://www.telegraph.co.uk/technology/video-games/6852383/Video-games-bigger-than-film.html*

Wiedenbauer, G. and Jansen-Osmann, P. (2008). Manual training of mental rotation in children, *Learning and Instruction* **18**(1): 30–41.

Wiedenbauer, G., Schmid, J. and Jansen-Osmann, P. (2007). Manual training of mental rotation, *European Journal of Cognitive Psychology* **19**(1): 17–36.

Appendix A - Data analysis tables and figures

	Statistic	df.	Sig
Post-Test RT	.965	40	.256
Control	.947	14	.519
Modified	.942	13	.485
Standard	.943	13	.497

Table 11: Testing normality assumption for the dependent variable

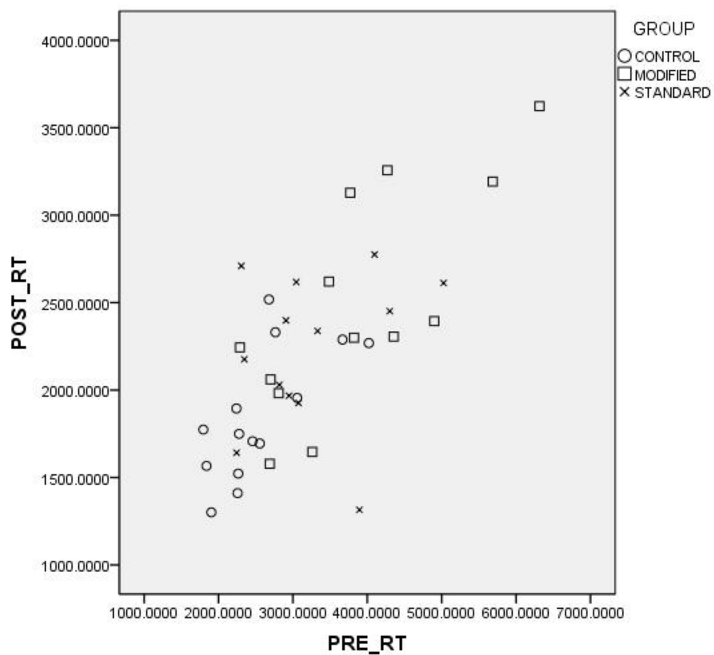

Figure 13: Scatter-plot of pre-test versus post-test RT with group markers (from SPSS)

	Statistic	df.	Sig
Control	.957	14	.672
Modified	.935	13	.400
Standard	.958	13	.726

Table 12: Testing normality assumption for difference scores

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	6096707.474	3	2032235.825	12.961	.000	.519
Intercept	3681686.405	1	3681686.405	23.480	.000	.395
pre_resp_time	3370924.826	1	3370924.826	21.499	.000	.374
group	218006.984	2	109003.492	.695	.506	.037
Error	5644726.595	16	156797.961			
Total	2.022E8	40				
Corrected Total	11741434.070	39				

Table 13: Tests of between subject effects – dependent variable POST_RT.

Change RT

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3178058.384	2	1589029.192	2.922	.066
Within Groups	20118910.552	37	543754.339		
Total	23296968.936	39			

Table 14: One way ANOVA on RT difference scores

		FIRST_10_LINE	LAST_10_LINE	FIRST_10_T	LAST_10_T	FIRST_10_Z	FIRST_10_Z_SANS_OUTLIER	LAST_10_Z	FIRST_10_L	LAST_10_L	FIRST_10_SQUARE	LAST_10_SQUARE	FIRST_10_SCORE	LAST_10_SCORE	PRE_RESP_TIME	POST_RESP_TIME
FIRST_10_LINE	Pearson Correlation	1	.409	.506	.268	-.055	.142	-.192	.515	.496	.033	-.101	.277	.341	-.089	.039
	Sig. (2-tailed)		.166	.078	.377	.858	.659	.529	.072	.085	.915	.742	.360	.254	.772	.899
LAST_10_LINE	Pearson Correlation	.409	1	.374	.486	-.141	.194	-.017	.302	.568*	-.262	-.286	.596*	.455	-.387	-.265
	Sig. (2-tailed)	.166		.208	.093	.646	.545	.956	.315	.043	.388	.344	.032	.119	.191	.381
FIRST_10_T	Pearson Correlation	.506	.374	1	.621*	.434	.475	-.023	.754**	.653*	-.035	-.233	.588*	.347	-.556*	-.369
	Sig. (2-tailed)	.078	.208		.024	.138	.119	.940	.003	.016	.910	.444	.035	.245	.048	.214
LAST_10_T	Pearson Correlation	.268	.486	.621*	1	.086	.456	-.027	.509	.892**	-.127	-.143	.382	.389	-.279	-.021
	Sig. (2-tailed)	.377	.093	.024		.779	.136	.929	.075	.000	.680	.641	.198	.188	.355	.945
FIRST_10_Z	Pearson Correlation	-.055	-.141	.434	.086	1	1.000**	.602*	.453	.114	.626*	-.038	-.092	-.080	-.177	-.439
	Sig. (2-tailed)	.858	.646	.138	.779		.000	.030	.120	.711	.022	.902	.766	.795	.563	.133
FIRST_10_Z_SANS_OUTLIER	Pearson Correlation	.142	.194	.475	.456	1.000**	1	.448	.574	.510	.304	.013	.318	.195	-.084	-.686*
	Sig. (2-tailed)	.659	.545	.119	.136	.000		.144	.051	.090	.336	.969	.313	.543	.795	.014
LAST_10_Z	Pearson Correlation	-.192	-.017	-.023	-.027	.602*	.448	1	.114	.134	.432	-.032	-.209	.115	-.245	-.255
	Sig. (2-tailed)	.529	.956	.940	.929	.030	.144		.710	.662	.141	.918	.493	.707	.420	.400
FIRST_10_L	Pearson Correlation	.515	.302	.754**	.509	.453	.574	.114	1	.664*	-.107	-.482	.451	.001	-.303	-.212
	Sig. (2-tailed)	.072	.315	.003	.075	.120	.051	.710		.013	.727	.095	.122	.997	.315	.486
LAST_10_L	Pearson Correlation	.496	.568*	.653*	.892**	.114	.510	.134	.664*	1	-.175	-.247	.331	.298	-.460	-.069
	Sig. (2-tailed)	.085	.043	.016	.000	.711	.090	.662	.013		.568	.417	.269	.322	.114	.824
FIRST_10_SQUARE	Pearson Correlation	.033	-.262	-.035	-.127	.626*	.304	.432	-.107	-.175	1	.510	-.469	.054	.354	-.168
	Sig. (2-tailed)	.915	.388	.910	.680	.022	.336	.141	.727	.568		.075	.106	.860	.236	.584
LAST_10_SQUARE	Pearson Correlation	-.101	-.286	-.233	-.143	-.038	.013	-.032	-.482	-.247	.510	1	-.320	.029	.186	.040
	Sig. (2-tailed)	.742	.344	.444	.641	.902	.969	.918	.095	.417	.075		.286	.925	.543	.896
FIRST_10_SCORE	Pearson Correlation	.277	.596*	.588*	.382	-.092	.318	-.209	.451	.331	-.469	-.320	1	.563*	-.331	-.389
	Sig. (2-tailed)	.360	.032	.035	.198	.766	.313	.493	.122	.269	.106	.286		.045	.269	.188
LAST_10_SCORE	Pearson Correlation	.341	.455	.347	.389	-.080	.195	.115	.001	.298	.054	.029	.563*	1	-.081	-.206
	Sig. (2-tailed)	.254	.119	.245	.188	.795	.543	.707	.997	.322	.860	.925	.045		.791	.500
PRE_RESP_TIME	Pearson Correlation	-.089	-.387	-.556*	-.279	-.177	-.084	-.245	-.303	-.460	.354	.186	-.331	-.081	1	.239
	Sig. (2-tailed)	.772	.191	.048	.355	.563	.795	.420	.315	.114	.236	.543	.269	.791		.431
POST_RESP_TIME	Pearson Correlation	.039	-.265	-.369	-.021	-.439	-.686*	-.255	-.212	-.069	-.168	.040	-.389	-.206	.239	1
	Sig. (2-tailed)	.899	.381	.214	.945	.133	.014	.400	.486	.824	.584	.896	.188	.500	.431	

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Table 15: Correlation matrix for research question 3 (from SPSS)

	Shapiro-Wilk		
	Statistic	df	Sig.
FIRST_10_Z	.752	13	.002
LAST_10_Z	.970	13	.899
FIRST_10_LINE	.955	13	.677
LAST_10_LINE	.942	13	.487
FIRST_10_T	.905	13	.159
LAST_10_T	.897	13	.121
FIRST_10_L	.897	13	.123
LAST_10_L	.968	13	.864
DIFF_T_SHAPE	.965	13	.823
DIFF_Z_SHAPE	.887	13	.088
DIFF_L_SHAPE	.932	13	.366
DIFF_LINE_SHAPE	.964	13	.808
FIRST_10_Z_SANS_OUTLIER	.945	12	.561

Table 16: Testing the assumption of normality for research question 2

Appendix B - Informed consent form

Researcher: Blaize Kaye
Masters (Cognitive Science) student
blaize.kaye@gmail.com

Supervisor: David Spurrett
Philosophy and Ethics UKZN
spurrett@ukzn.ac.za
Memorial Tower Building, Howard College Campus

Purpose of the Study: The purpose of this study is to investigate what effect playing the fast paced puzzle game Tetris has on Mental Rotation performance.

Procedures to be followed: You will complete a mental rotation test. You will then be randomly assigned to one of three groups. If you have been assigned to one of the two experimental groups you will play Tetris once a week for six weeks. If you are assigned to the control group you will meet with the researcher once a week to complete a question based task and video-game. At the end of the six week training period you will complete a second Mental Rotation test.

Duration/Time: Each Mental Rotation test will last approximately an hour. Participants in the experimental groups will undergo, minimally, five hours of Tetris training, while control group participants will spend at least five hours on their question based tasks and video-game.

Risks/Discomforts: There are no risks to you.

Benefits: By participating you will stand a chance of winning prizes to the value of R500.00

Statement of Confidentiality: Your participation in this research is confidential. No personally identifiable information will be reported or published.

Access to Biographical information: Please be advised that your involvement in this research requires your consent for the release of information from the university's computer system. None of this information will be reported or published and your anonymity will be protected at all times.

Right to Ask Questions: Any questions about the study can be asked via email to the researcher.

Voluntary Participation: Your participation in this research is voluntary. You can withdraw at any time. Refusal to take part in or withdrawing from this study will not prejudice you in any way. You must be 18 years of age or older to take part in this research study.

I .... (full names of participant) hereby confirm that I understand the contents of this document and the nature of the research project, and I consent to participating in the research project. I understand that I am at liberty to withdraw from the project at any time, should I so desire.

_________________ _________________

Signature Date

Appendix C - Text advertising research – used for both posters and electronic notices

Are you interested in winning prizes for playing computer games?

A few of us up in MTB are looking into the psychology of computer game playing and would love to have you take part in our study.

When? : The study will run over 6 weeks through the second semester.

What times? : We have slots available on Tuesdays, Wednesdays, and Thursdays (all starting at 2PM).

How much time will it take? : It shouldn't take more than an hour a week

What do I get ? : We have daily prizes, to the value of R500, up for grabs.

How do I apply? : You can browse to http://gamesforscience.co.za, sign up, and we'll contact you.

Other questions? : If you have any more questions you can send an email to gamesforscience@gmail.com

This research has been granted full approval by the UKZN Research Ethics Committee

Appendix D - Briefing text

Welcome text - read to all groups at the first session.

Firstly, I would like to thank you for your willingness to participate in our research. I'm sure that our time together will be fun and productive.

During the course of this, our first session, we will begin with a short Mental Rotation assessment - a task I'll explain in some detail in a moment. Once that is complete we will be introducing you to the game you will be playing, and you will spend what's left of today's session playing this game.

We will meet five more times after today. At each of these sessions, except the last, you will be engaged in your assigned tasks. Our final session will comprise solely of a second Mental Rotation assessment, and some questionnaires.

Please note that we will not be meeting during the mid-semester break, this is the week of the 26th to the 30th of September.

If there is any reason why you will not be able to make a particular session, please get in contact with me and I will do my best to make a plan to accommodate you. Obviously it is best if everyone attends all their scheduled sessions, but I understand that this is sometimes not possible. I would much rather we try and make a plan that will work than for you not to attend a session, or feel that you need to drop out from the study. Again, do not hesitate to contact me.

Every week there will be a lucky draw, where one participant will receive a prize and, just by being present at a session you will be eligible to win. At the end of the session we will run a program that will randomly select a student number from the list of people who have logged into the system. Once the computer program has run and outputted its results, we will announce who that week's lucky winner is, immediately after which they will receive their prize.

Instructions read to all groups before pre- and post-tests of MR performance.

You will now partake in a standard assessment of your Mental Rotation ability. In this assessment, you will be presented with two images placed side by side. Your task is to identify, as quickly but also as accurately as possible whether the image on the right hand side of the screen is only a rotated version of the image on the left or whether, in addition to being rotated, the right hand image has also been reflected.

I will now explain the difference between when the images have been only rotated or rotated as well as reflected.

When the image on the right hand side has been only rotated we mean that if you took the image on the right and spun it around its centre, you could get it to look exactly like the image on the left without needing to do anything else.

<demonstration here – take the identical image board and spin the right hand one around its centre>

When we say that the image has also been reflected we mean that in addition to spinning the image around on its centre, we would need to flip it over to make it look like the image on the left.

<demonstration here – take the reflected image, rotate it to the same orientation as the image on the left, and show the reflection by physically flipping the image around>

It is very important that everyone is 100% comfortable with the differences between when an image is only rotated when it is rotated and reflected. I am more than happy to go through the explanations and demonstrations again if anyone feels that they would like, or need, some further clarification.

Is there anyone who feels that the demonstration was not completely clear, or who feels that they don't fully understand the differences?

<pause for any requests to go through demonstration again>

During the task you will not be required to physically rotate or flip any images, rather, given the two images, you are tasked with imagining, or visualizing, whether you would need to only rotate the image on the right to make it look identical to the one on the left, or, in addition to rotating the image, whether you would have to also flip it to get them to look identical.

Underneath the two images you will find two buttons – the one on the left is labeled "rotated only" the one on the right is labeled "reflected". If you think that the image on the right need to only be rotated to get it to look identical to the image on the left, you will click the button on the left labeled "rotated only".

If you think that it also needs to be reflected, you will press the button on the right hand side labeled "reflected". As I mentioned above, we are interested in both the speed and accuracy of your responses, so it is important for you to try and answer as quickly as you can, but not so quickly that you make mistakes.

Does anyone have any questions about what is required of you for each pair of images?

<pause here for questions>

The task will proceed as follows. The first screen you will see once the system has loaded will ask you to enter your Full name and Student number – after accurately filling these values in, you will proceed to a briefing screen that will, once again, explain what is required of you. Once you have moved on from this briefing screen, there will be 10 practice sets of images. Please use these 10 practice pairs of images to get used to the task and the interface – we will not record your responses on these 10 pairs. Once you have reached the end of your practice round there will be a screen telling you that you are about to move on to the real task. As soon as you click the button on this page to continue, you will have entered the live task environment where all of your responses will be timed and recorded. You should now be trying to answer as quickly and accurately as possible.

The images in the live task are not the same as those you will see in the practice round.

There are 160 pairs of images in total. Once you have compared all 160 pairs of images you will be presented with a screen telling you that "your responses will now be sent to the server". You will then press the "Continue" button and sit back from your computer while it communicates with the server. It takes about 30 seconds for all the data to be transferred so please do not touch anything during this time, if you navigate away from the page your data will be lost.

Are there any questions about the task?

<pause here for questions>

Thank you, please open a copy of Firefox and browse to the webpage

http://mr.enactlabs.com/bomoko/mental-rotation.html

If the browser asks if this is a trusted host, click yes, or – if you need help – call me over and I will get it working.

Explanation of Training task read to Standard and Modified groups after pre-test had been administered.

I will now describe the task that you will be engaged in for the next six weeks.

Each of you have been randomly assigned to play a different version of the fast-paced puzzle game "Tetris". For those of you who are unfamiliar with the game, the aim of Tetris is to prevent an ever growing wall of shapes from reaching the top of the game area. The wall is built up from shapes, called "zoids", that fall from the top of the screen. As these zoids fall you control the way in which they fall by using the arrow keys on your keyboard. Pressing the "up" button rotates the zoid. Pressing the "left" and "right" buttons moves the falling zoid left and right. Finally – if you are satisfied with how the zoid is orientated and positioned, you can press the "space" button and the zoid will immediately fall to the bottom. In order to prevent these zoids from building a wall that reaches to the top of the game area you need to form completely filled horizontal rows. Every complete row you form will disappear, reducing the overall height of the wall.

While playing the game, please try to keep the following in mind. We are interested in your problem solving skills, as such it is important to play as well as you possibly can. You will find that you will improve steadily as the study progresses, so try your best.

Does anyone have any questions related to Tetris?

<Pause now and take any questions related to Tetris>

Thank you. I would now like you to open a copy of Firefox and browse to the webpage http://tetris.gamesforscience.co.za Use your student number as your user name and password. Once you are logged into the system, click the item "Play Tetris" on the menu along the left hand side of your screen and our version of Tetris will open. Click "New Game" to start playing.

Explanation of Training task read to Control group after pre-test had been administered.

I will now describe the tasks that you will be engaged in for the next six weeks.

There are two parts to your task.

Firstly - Each of your sessions will begin with a series of questions in which you will be asked to state your preference between two amounts of money. One of of the amounts will be available immediately, while the other will be available only after a certain delay.

For example, you may be asked whether you'd prefer R500 right now, or R1000 in 80 days time. Each choice will require you to select the check box next to either the immediate or delayed option. Once you have checked one of the options you will then click on the button labeled "Confirm Choice". This will confirm your choice, and take you to the next question. There are 100 of these choices each session.

While answering these questions we would like you to keep the following in mind. First, please try to answer the questions seriously, as if you were actually going to be receiving the money. Second, consider each choice separately, try not to make your choice based on the choices that have come before or expect to see on the next choice. Finally, remember that there are no right or wrong answers, we are interested in how you would choose given the choice that you are offered.

Are there any questions about this first part of the task?

<Pause now and take any questions related to part 1>

Once you have made all 100 of your choices the system will give you a link to the second part of your task, which is on online version of the popular puzzle game "Lemmings". The object of this game is to guide as many of these animated characters, called "Lemmings", as you can to the designated exit point for each stage of the game. This is accomplished by assigning certain of the "Lemmings" one of eight different skills that allow the "Lemming" in question to either navigate past obstacles or change the landscape of the stage – through building, digging, and demolishing walls - in such a way that the other "Lemmings" will be able to walk to the level's exit. You will find an explanation of each of the various skills that you can assign to your "Lemmings" on the handout. Each level is a bit easier than the one after it, so the game gets more challenging. Try to get as far as you can - we are interested in the development of your problem-solving skills..

Each of you will receive a “Lemmings log” each session. Please fill in your Name, Student number, and the date on the log. You will see a list of all “Lemmings” levels, for each of the four difficulty ratings (fun, tricky, taxing, and mayhem) and next to each you will see the code that will unlock each of those levels. You’ll find that your skill with the game will develop best if you begin with the easiest levels and work your way up through the levels as they get more difficult. The codes will let you pick up from where you left off each session.

Once you have finished a particular level, write down the percentage of “Lemmings” you managed to save on your attempt.

<Show page and Demonstrate>

Are there any questions about this second part of the task?

<Pause now and take any questions related to part 2>

Thank you. I would now like you to open a copy of Firefox and browse to the webpage http://tetris.enactlabs.com Use your student number as your username and password. Once you are logged into the system, click the item "Begin Part 1" on the menu along the left hand side of your screen and the first part of your task will begin.

Appendix E - Mental Rotation test instructions

Are the images the same, or different?

In what follows you will be presented with a series of image pairs. The aim of this exercise is to tell us, as *fast and as accurately* as you possibly can whether the images you have been presented with are the only rotated, or also reflected.

All of the pairs of images *look* as if they might be the same, but many of them are actually mirror images of each other (and therefore different - reflected).

If you're see that the images are mirror images of each other, you are required to click the button labeled "Reflected", if they are exactly the same image you'll click the button labeled "Rotated".

You will be presented with a short test round consisting of 10 items - once the test round is done, the assessment will begin.