Advanced Research and Design

Designing Educational Software To Improve Cognitive Abilities; Pilot Study Results

This report summarizes the preliminary findings of a ten-week pilot study evaluating the efficacy of a software-based visual-spatial intervention. Between June 17th and August 31st, twenty-two children with a variety of learning difficulties participated in thirty (30) two-hour sessions using specially designed cognitive development software. Visual-spatial neuropsychological tests were administered before and after the ten-week pilot program. The data analysis yielded three key findings:

1) On average the sample's visual-spatial scores improved;

2) Students diagnosed with a learning disorder but not with attention deficit disorder improved by a greater degree than students diagnosed with both a learning disorder and an attention disorder

3) Achievement on the visual-spatial software is correlated with performance on neuropsychological tests of visual-spatial skills.

Paper Presented at
Interactive Technologies Conference on Training,
Education, and Job Performance Improvement
Arlington. Virginia
August 6-8, 2003

David A. Stevens • Michael W. Cornell • Sharon Colvin

Paul Schwarz • Roy Pardi • Beth Pilgrim

Introduction

The current trend in efforts to improve educational performance is to raise academic standards, support them with increased instructional objectives, and follow up with regular assessments. The research[1] summarized in this document pursues an alternative path to improving educational performance - the cognitive intervention. Here we use the term 'cognitive intervention' to mean intervention in the process of cognitive development. This is achieved by providing students with the appropriate level of cognitive challenge to foster higher levels of thinking. The goal of these interventions is the development and improvement of fundamental cognitive abilities that will transfer to more efficient and meaningful school learning.

The topic of cognitive interventions is not an uncontroversial one. On one side of the debate, a number of researchers point to considerable evidence that cognitive abilities and school achievement can be improved through sophisticated cognitive developmental interventions (e.g., Adey & Shayer, 1994; Case & Griffin 1993; Stevens 2000). However, despite the evidence that they can in some cases produce impressive gains in intellectual abilities and academic performance, these interventions have only been carried out on small scales and are rarely replicated. As a result, they have been unable to effectively challenge the conventional wisdom that cognitive abilities are fixed from birth, and there are many, people who maintain that improving cognitive ability through special interventions is for all intents and purposes an unachievable goal (e.g., Detterman, 1993; Herrnstein & Murray, 1994).

Given that examples of effective cognitive interventions do exist, it follows that cognitive abilities can be improved under the right conditions. The really important question, therefore, is whether these abilities can be improved on a wide enough scale and in an efficient enough manner to make this approach a feasible solution to poor school achievement. The central issues here are implementation and scalability of the intervention.

The purpose of this pilot study is to determine if technology can be leveraged to make cognitive interventions easier to implement and more widely available. The research plan is to design software-based cognitive development programs based on examples of the successful classroom programs already documented and then test the software with students to determine if its use is associated with measurable gains on tests of cognitive ability

[1] The research described in this paper is supported by grant #2000-00-4002a from the Advanced Technology Program of the National Institute for Standards and Technology. The data presented, statements made, and the views expressed are solely the responsibility of the authors.

Rationale for the Project: Improving Cognitive Abilities

Research has shown that cognitive skills can be developed in the classroom or clinical setting through intervention programs, and that the resulting improvements are associated with higher academic performance (for an in-depth review see Stevens, 1999).

Robbie Case and his team of researchers at the University of Toronto initiated the *RightStart* program after recognizing that many children fare poorly in math because they lack the prerequisite cognitive skills. Their hypothesis is that providing the conceptual foundation through a cognitive intervention will transfer to improved math performance on a broad array of tasks. Their results indicate that the newly acquired conceptual understanding transfers to improved performance on post-test measures of quantitative ability not covered in the intervention. The acquired knowledge also appeared to transfer to superior mastery of later mathematics material when compared to control groups.

The authors draw four conclusions from the results of their study. First, children can develop the key underlying cognitive structure through the *RightStart* cognitive intervention. Second, the *RightStart* program outperformed two control programs in developing this knowledge. Third, the development of the cognitive structure improved the subjects' performance on related domains that the training did not specifically cover. Fourth, the development of this conceptual structure leads to improved learning in later educational settings. (Griffin et al., 1995)

Philip Adey, Michael Shayer and colleagues offer another example of improving academic performance through a cognitive intervention program. They created an intervention to develop general cognitive processes in the context of a traditional science curriculum (*Cognitive Acceleration through Science Education).* This approach is inspired by empirical insights generated by the systematic measurement of the cognitive development of large populations of school children (Adey and Shayer, 1994). The authors discovered that the average student does not progress through the stages of cognitive development nearly as fast as is normally quoted in the cognitive science literature, and that there is a far wider spread of cognitive ability than had been previously reported.

The *CASE* program is intended to develop the 'general cognitive processor' of each student to enable more efficient and complex academic learning. Post-tests administered immediately after the intervention revealed improvement in cognitive ability.

Achievement tests administered two and three years after the completion of the intervention indicated gains in English, Science, and Mathematics when compared to a control group. Adey and Shayer (1994) write that these results are "very encouraging to those who reject the idea of intelligence as a fixed potential and believe that educational intervention rooted in well-established theories of cognitive development can have long-term and replicable effects on young adolescents' academic achievement (p. 112)."

Kvashchev carried out an experiment in former Yugoslavia to determine if he could develop 'creative problem solving' through an educational intervention with high school students. The Creative Problem Solving program consisted of a series of exercises designed to develop fluid intelligence (Gf). Kvashchev hypothesized that 'general cognitive schemas' could be developed through a training program if it was long enough in its duration, frequent enough in its occurrence and if he grounded its theoretical basis in psychological research (Stankov, 1986, p. 210). Stankov and Chen's analysis of the Creative Problem Solving data concluded that both fluid and crystallized intelligence increase as a result of general cognitive abilities training. "We have to conclude that training in creative problem solving is effective, that transfer occurs and, above all, that performance on tests of general intelligence can be improved (Stankov, 1991).'"

Furth and Wachs (1971) outline an educational intervention for elementary school children in their book *Thinking Goes to School.* The goal of this program is to provide students with the prerequisite cognitive skills to be able to thrive in an intellectually challenging environment. The curriculum focuses on developing skills such as logical reasoning, visual thinking, and sensorimotor coordination. An eight-year longitudinal study of the program revealed that, relative to a control group, the children who participated in the intervention scored higher on academic achievement and cognitive ability tests year after year through seventh grade (Stevens, 2000).

There are also cognitive intervention programs in individualized clinical settings that have yielded positive outcomes relative to control groups. The Structure of Intellect (Guilford, 1967; Meeker, 1969), Instrumental Enrichment (Feuerstein, 1980; Feuerstein et al, 1981), and the Perceptual Enrichment Program (Theisen, 1998) are a few examples.

The mainstream approach to improving academic achievement involves frequent testing of the defined curricula and teaching specifically to those curricula or even teaching to the test. Studies on the efficacy of cognitive intervention programs demonstrate that these interventions offer a valid complementary approach. Rather than teaching to the academic tests, teachers can improve academic performance by developing the broad underlying cognitive faculties that support students' learning in the more specific academic domains.

Difficulties with Implementation

These interventions offer an optimistic view of the plasticity of cognitive skills and present a complementary approach for improving school learning. Unfortunately none of these programs have been widely or systematically implemented, replicated, or evaluated in the US. It is possible that in each case the complexity of the interventions are such that they are not easy to master and implement by those who are unfamiliar with the theoretical foundation. For example, the *Thinking Goes to School* program of Furth and Wachs has been identified as a model program in the state of Pennsylvania based on the success of an elementary school district that installed the program in 1990. Despite frequent visits by other school administrators in the state, the program has not been implemented beyond this initial district.

Most of the successful cognitive intervention programs have been implemented—and in some cases staffed —by the program designers, who are expert in the theory and techniques. These programs are often complex curricula based on a dramatically different conceptualization of education, its relationship to the learner, and the role of the teacher.

For example, the *RightStart* program has achieved impressive results with children from families of low socio-economic status (Griffin, Case, and Siegler, 1994). However, the team of researchers who implemented the program caution that it only can be effectively replicated by a staff with doctoral level training (Kalchman, 2001).

Some of the other barriers to widespread implementation include the amount of time needed to train staff, the need (in some cases) for a lower student-teacher ratio, and the perhaps counter-intuitive assumption that focusing on cognitive development can improve academic performance. The research on cognitive interventions supports these programs potential to improve school learning. However, there are considerable barriers to their widespread implementation and evaluation.

A Possible Solution: Using Technology to Overcome Barriers to Implementation

One of the main obstacles to the widespread implementation of cognitive development interventions is their need for a lower teacher-to-child ratio. This can make the implementation of such interventions cost prohibitive for most schools. Another key obstacle is the depth of knowledge needed by the teachers in order to effectively implement the programs. The curricula and techniques of cognitive developmental interventions are not common in most teacher training programs and often require a different teacher role than in traditional academic settings.

Educational Software Based on the Principles of Successful Cognitive Interventions

In order to design educational software to develop cognitive abilities, we analyzed the key components of effective programs and extracted the common principles and techniques that were fundamental to their success. We found that having key strategies generated from cognitive science frameworks set these programs apart from more traditional educational practices.

The most striking similarity is that nearly all of these programs draw upon the Piagetian theoretical model of cognitive development (Stevens, 1999) -and upon two unifying principles in particular. First is the notion of powerful organizing schemata that underlie thought and action. Problem-solving and learning ability is not determined only by processing speed and accumulated cultural knowledge, but that the actual structure of the mind lends itself to more or less complex and effective forms of understanding. Second is the notion that individuals construct intelligence through interaction with the environment. It is by acting upon the environment in a meaningful way and integrating feedback to modify our understanding that we develop more complex thinking abilities.

Although usually implicit, another element common among many of the successful cognitive interventions is a reliance on the idea of developmental range (related to Vygotsky's ''zone of proximal development"). That is, they recognize that in order for an educational interaction to be successful, it needs to fall within the child's range of developmental capabilities. Other commonalities include a reliance on problem solving and active participation. None of the programs described above uses direct teaching of answers or strategies. Instead, they present developmentally appropriate challenges that the student needs to resolve. Each program builds feedback for the student into each challenge.

Another commonality is that most of these studies demand considerable intervention time both in terms of intensity and overall duration. This indicates that the process of transforming cognitive skills is not a short-term endeavour, but needs to be a regular part of a child's schooling over an extended period of time.

These common principles of effective cognitive interventions are examples of some of the guiding pedagogic principles that have informed the design of Lexia's software-based cognitive developmental interventions.

Developing Visual-Spatial Skills

In the 1980's a number of cognitive theories emerged that offered a view of the human mind as being organized around several key domains fundamental to school in particular and overall life adaptation in general (e.g. Gardner, Fodor, Sternberg, Case). While the prevailing notion was that individuals "possessed" a general intellectual level of functioning (e.g. **I.Q.),** these theories argued that individuals had a variety of domain specific abilities that supported school learning. Some of the different theorists break down these domains along similar lines (although they might call them by somewhat different names).

Following the lead of these theorists and the designers of cognitive intervention programs who also focus on specific domain-centred abilities. We selected to address the following cognitive domains with the understanding that they are fundamental to school learning and achievement:

Visual-Spatial Ability
Logical Reasoning Ability
Expressive Communication Ability
Receptive Communication Ability
Auditory Imaging Ability

The project began in the domain of visual-spatial abilities. We chose this area as it appeared to be the most neglected by classroom curriculum designers as well as educational software designers. Visual-spatial skills have been consistently linked to a host of low- and high-level skills that are central to academic and even occupational performance but are not directly addressed in most traditional school programs (cf., Haywood and Coltheart, 2000; McCormack, 1988). Moreover, it is likely that even greater demands will be placed upon an individual's visual-spatial abilities as we increase our dependence upon computers for communication, data visualization, simulation, education, and generally for interfacing with machines in many facets of life. Researchers have signalled that we are in the midst of developing a reliance on technologies such as the internet and computer simulations that require well developed visual-spatial abilities (e.g. Gagnon, 1985; Compaine, 1983; Subrahmanyam and Greenfield, 1992).

There is also evidence that visual-spatial skills can be improved through intervention programs. Several studies provide evidence that visual-spatial intervention programs can improve problem-solving on visual-spatial tests, and performance on these tests correlates with achievement in academic subjects and occupational skills (cf, Moses, 1979; Eastman and Carry, 1975; Hill and Oberhauf, 1979; Lowery and Knirk, 1982-83; Gagnon, 1985; Subrahmanyam and Greenfield, 1992; McClurg, 1992). Although the link from visual-spatial intervention to academic and occupational performance delineated by these studies is at present indirect, it is certainly promising enough to warrant additional investigation. In the following sections we describe our efforts to study it further.

The First Module; Designing Software to Develop Visual-Spatial Skills.
Before a visual-spatial software curriculum could be developed, we needed to determine if visual-spatial tasks could be administered on a desktop-computer typically found in schools. Our first question was whether a digital simulation of visual-spatial learning activities were comparable to the real-world tasks found in programs such as *Thinking Goes To School, Instrumental Enrichment, Structure of Intellect* and the *Perceptual Enrichment Program.* The digital computer would merely simulate the hands-on manipulatives and other tactile media that are the focus of the 'real-world' interventions. We were concerned that this simulation could alter the experience in such a way that the computer-based tasks would not be addressing visual-spatial skill but some other ability (most likely logic). It was also possible that the computer-based version would be excessively difficult or easy thereby providing too much or too little cognitive challenge. Any of these issues could easily undermine the effectiveness of a computer-based program and therefore necessitated empirical testing.

In order to answer this key question, we designed the following study (for an in-depth review see Stevens et. al. 2002). We adapted a series of eleven progressively complex visual-spatial tasks from a visual-spatial intervention program that had demonstrated efficacy (Furth and Wachs, 1971; Wachs, 2000; Stevens, 2000). These tasks consist of matching or transposing designs of three wooden parquetry blocks (square, triangle, diamond). We developed a software program consisting of eleven virtual visual-spatial tasks that closely match the appearance and functionality of the 'real world' tasks. Third, we developed a novel maze navigation software intervention program to develop visual-spatial skills that consisted of seven visual-spatial tasks. This consists of a three-dimensional maze environment where the student needs to navigate the space according to a small two-dimensional map. The hierarchy of tasks that comprise these activities are designed to be of increasing complexity and organized in such a way that successive tasks call for increasingly complex visual-spatial skills.

A sample of 76 children between the ages of six and ten years of age participated in this study. Each child was given the series of tasks from the three activities (''real world blocks', "virtual blocks", "virtual maze") until she successfully completed all tasks in an activity or until she failed three consecutive tasks within an activity.

Rasch analysis and two-sample t-tests were used to analyze the data and determine if the visual-spatial tasks from the "real" activity and the "virtual" activities fell along the same cognitive dimension. From our data analysis we concluded that the computer-based and the real-world visual-spatial tasks tap a single underlying ability. We found that 28 of the 29 tasks lined up within the margin for error along a single dimension. From this we inferred that our computer-based interventions were indeed addressing similar visual-spatial skills as the real-world interventions to which we compared them.

The second question we addressed in this study was whether the computer-based visual-spatial tasks were of the same difficulty level and followed the same hierarchical progression as the real-world activities. We found that on average the performance of the seventy-six children on the computer-based tasks did not differ significantly (at the 0.05 level) from their performance on the real-world tasks and that the hierarchical order of difficulty was consistent.

From this study we determined that visual-spatial tasks could be created for the desktop-computer that addressed the same cognitive abilities as real-world tasks and that their difficulty levels were comparable. These findings supported the scientific rational of our project and enabled us to go forward with our creation of visual-spatial educational software programs

Defining Visual-Spatial Skills

The domain of visual-spatial skills has been widely studied, but there are few systematic and unifying formulations that define the different visual-spatial abilities. There is a considerable degree of variability in what researchers mean when they refer to visual-spatial skills. From our review of this literature we identified 22 distinct visual-spatial sub-skills.

Table 1. The 22 visual-spatial skills addressed by the software.

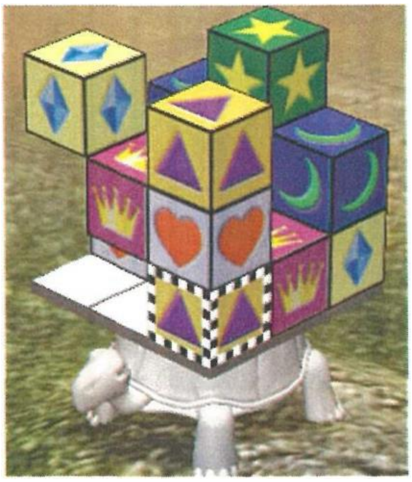
2D-3D COORDINATION	VISUALIZATION
MULTI-PESPECTIVE COORDINATION	VISUAL TRACKING
FIGURE GROUND	VISUAL PURSUIT
VISUAL DISCRIMINATION	SPATIAL LOCALIZATION
ABSTRACT MATCHING	SPATIAL RELATIONS
CONCRETE M ATCHING	SPATIAL ORIENTATION
VISUAL FIND	PERSPECTIVE TAKING
VISUAL MEMORY	REFLECTIONS
MENTAL ROTATIONS	CONCRETE PART TO WHOLE RELATIONS
MULTIPLE MENTAL ROTATIONS	ABSTRACT PART TO WHOLE RELATIONS
REPRESENTATIONAL THOUGHT	NEGATIVE SPACE

Five Software Activities to Improve Visual-Spatial Abilities

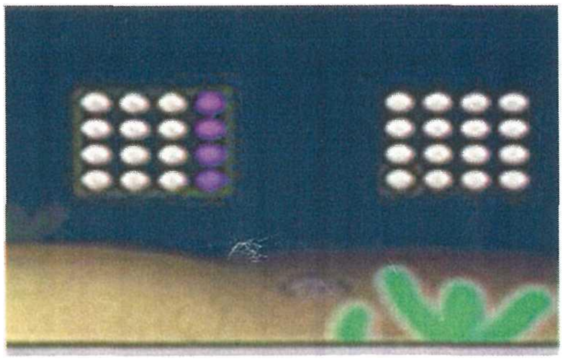
We developed five software activities to improve visual-spatial abilities. In order to ensure the activities were suitable for children we adapted a game controller typical of many home video game systems. We found that the game controller was relatively inexpensive and better suited for children than a mouse and keyboard. During the research and development process we conducted systematic user-interface testing and cognitive content testing with more than one hundred children. This testing informed both the development of interactive three-dimensional graphical user interfaces as well as the order and complexity of the visual-spatial tasks. We employed statistical techniques to objectively confirm the appropriate hierarchical order of the tasks and to ensure that the activities addressed the same skills as the proven clinical and classroom intervention programs, as described above.

We developed five activities to exercise the twenty-two visual spatial skills previously identified. We created an interface that enables children to choose when they visit each activity and to track their progress within each activity.

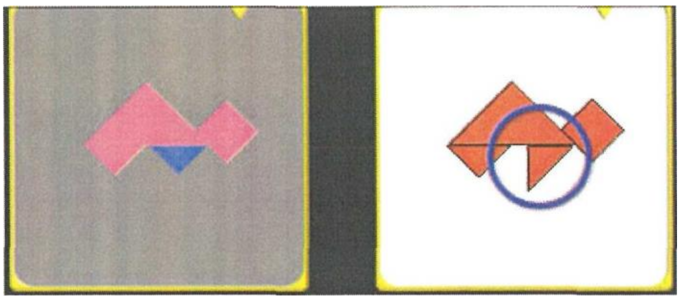
Spatial Delivery is a three-dimensional navigation activity that challenges children to use a map to find their way through a maze to deliver a package. *Spatial Delivery* emphasizes the development of navigation, spatial directionality, spatial memory and coordination between two-dimensional information and a three-dimensional environment. Unlike many computer games, which mostly rely only on visual memory and reaction speed, Spatial Delivery has been designed to challenge children to develop progressively more sophisticated visual-spatial abilities.



Cubes was designed primarily to improve perspective taking. The activity consists of comparing a three-dimensional design of stacked cubes to a two-dimensional representation of those cubes from a variety of perspectives. At first a child needs to select the cube in the 3D design that is represented on a 2D card. At intermediate levels the child needs to colour in the cards to represent how the 3D cube design would look from a variety of perspectives. At the highest levels users need to construct the 3D cubes design based on 2D representations of it from a variety of perspectives (e.g., front, top, side). In all cases the child is challenged to coordinate multiple 2D representations with a 3D object.



Waterworld is a two dimensional activity that involves creating matching coloured grids. At higher levels, children need to create matching line patterns on the grids and eventually create designs that have been rotated 90 or 180 degrees. *Waterworld* was designed to improve figure-ground perception, understanding of negative space, visual matching and visual memory.



Tangrams is adapted from the ancient Chinese game of the same name. It was designed primarily to develop part-to-whole relations and visual transformations. Part-to-whole relations involves trying to create a complete design from its component parts. The activity begins with first identifying matching individual pieces and gradually progresses to the more commonly known Chinese Tangram activity. This computer-based version of Tangrams breaks down the steps to provide the necessary foundation abilities to eventually succeed at the harder Tangram puzzles. This activity also adds tasks that involve turning and matching the designs from memory.

The Sample

We recruited a sample of children to test the effectiveness of the software. We sought children who had known difficulties with visual-spatial abilities and associated learning challenges. We sent recruitment letters to educational professionals and schools in the Boston area. Given the extensive time commitment of the pilot study many families wereunable to participate in the program once they learned of the demands of the protocol.

We accepted twenty-three children to participate. One child was not accepted because of his severe difficulties with ocular-motor and hand coordination. One subject dropped out of the pilot study after seven days of participation because she was not able to meet the time demands of the protocol.

The sample consisted of twenty-two children between the ages of seven and eleven years (mean age = 9;9). Twenty children were Caucasian, one Native American and one African American. Sixteen of the children were male, six were female. One child came from a family with an annual income between $40,000 and $65,000. Ten children came from families with annual incomes between $65,000 and $99,000. The rest of the sample came from families with annual incomes above $99,000.

The demographic profile of the sample represented the middle to upper class suburbs that surround our research lab.

Parent reports revealed that twenty of the children had some type of learning disability and that nine of those children had been diagnosed with ADD or ADHD.

Administration of the Software Intervention

The twenty-two children came to the computer lab thirty times over the course of the summer to participate in two-hour sessions. The children took one scheduled twenty-minute break in the middle of the session, and took other breaks as needed during the session. The children were given prizes (value of $ 1 to $3) for attending each session At the end of the pilot study each participant was given a $50 gift certificate to a bookstore or a toy store. The prizes and gift certificates were not tied to the child's performance on the software. The children received their prizes for attending, and not for using the software.

At the beginning of the intervention there were two computer lab administrators supporting the children's use of the software at all times. This support came in the form of answering questions, providing instructions, support and encouragement. After the third week of the intervention there was usually only one administrator in the lab. There were not more than eight children in the lab at any one time.

Families scheduled the software sessions at their convenience over the course of the summer. The only constraint was that children could not participate in more than one session on the same day. Some children participated in sessions nearly every day for six consecutive weeks. Others spaced their sessions across the summer coming three or four times a week. One subject completed the intervention on the twenty-second session by working his way through all of the levels of the activities.

The Visual-Spatial Test Battery

Between one day and two weeks before the start of the software intervention each child was pre-tested with a battery of neuropsychological and cognitive tests that measure different visual-spatial abilities. Each child was post-tested with this same battery between one day and two weeks after the last software session. The entire battery took between sixty and ninety minutes to administer. The test battery consisted of the following seven measures:

Stanford-Binet Pattern Analysis:

The Pattern Analysis subtest challenges the subject to construct geometric designs from its component parts. The parts consist of cubes with a unique pattern on each side of the cube. The subject must identify the correct side of each cube to use and then assemble the cubes together to produce the indicated design. The patterns range in complexity from two cubes up to nine cubes. The subject is allowed between thirty and ninety seconds to construct the design, depending on the number of cubes involved.

Stanford-Binet Paper Folding:

The Paper Folding subtest asks subjects to visualize how a piece of paper that has been folded and cut into specific patterns will look after the paper has been unfolded. The subject is shown two practice tasks where a piece of paper is folded in half and a section is cut out. The subject is asked to pick from five possible answers which picture will look the same as the piece of paper when it is unfolded. For the actual test items paper is not used. The subjects are only shown a diagram of how the paper will be folded and how it will be cut. This test measures a child's ability to visualize how objects appearance changes when they have been rotated and manipulated in different ways.

Benton Test of Facial Recognition:

This test determines the ability of the child to distinguish similar from dissimilar faces. A target black and white photograph is presented to the child along with six apparently similar photographs. However, only one photograph is actually of the same person. As the test progresses the child must identify three of the six photographs as the same person at the target photograph. With each successive task the photographs have more shadows making it harder to distinguish the features.

Benton Test of Block Construction:

This test evaluates a child's ability to construct blocks that match a target photograph. There are three separate photographs with each one containing more blocks in the construction and a more complex organization. The blocks consist of cubes, rectangles and. rods of different lengths and heights.

Benton Judgment of Line Orientation:

The test of Line Orientation assesses a child's ability to distinguish lines drawn on similar or dissimilar angles. A target design is presented with two lines drawn 011 different angles. The child must compare this design to a design with all of the possible angles used in the test with each one labelled with a number. The child must report the numbers from the full spectrum of lines that correspond with the two lines in the target design.

Benton Test of Form Recognition:

This test evaluates a child's ability to identify matching designs. A target design is presented along with four somewhat similar designs. Each design has a subtle difference from the target design except for one. The child must identify the design that perfectly matches the target.

Test of Mental Rotations:

In the Test of Mental Rotations the child is given a four page worksheet that contains illustrations of the three-dimensional block designs. There is a target design on the left and four somewhat similar designs on the right. The child must identify the two designs on the right that are the exact same cube construction as the target but that have been rotated into different positions. The other two designs are not exactly the same cube construction.

Data Analysis

We conducted an analysis to determine if the post-test scores were statistically significantly higher than the pre-test scores. Using one-sided t-tests we determined that the post-test scores were greater to a statistically significant degree for four of the seven tests. In several of the tests we detected a ceiling effect where the test subjects were getting all or nearly all of the items correct on the post-test. This weakens the validity of the measure as it appears that in some cases the child's ability progressed beyond the measurement scale of the test. In some other cases a few children scored extremely well on the pre-test, not leaving much room to make progress on the post-test. *Pattern Analysis:* Table 2 summarizes the pre- and post-test scores for the Pattern Analysis test. On average the sample tested 6.09 points higher on the post-test than on the pre-test. A t-test was computed to test the hypothesis that the post-test score was not higher than the pre-test score (at the 0.05 level). The result of the hypothesis test ($t = 2.511$, $p < 0.006$) surpassed significance at the 0.05 level, which means the null hypothesis (post-test not greater than pre-test) can be rejected based on this sample. We take this as evidence that the sample performed better on the post-test than on the pre-test.

Table 2: Pre-, Post-, and Delta Scores for Pattern Analysis

	N	Mean	SD	Median	Min	Max	Range
Pre-Test	22	26.95	8.31	29.5	0	39	39
Post-Test	22	33.05	7.77	35.5	10	42	32
Delta	22	6.09	—	6	-	-	—

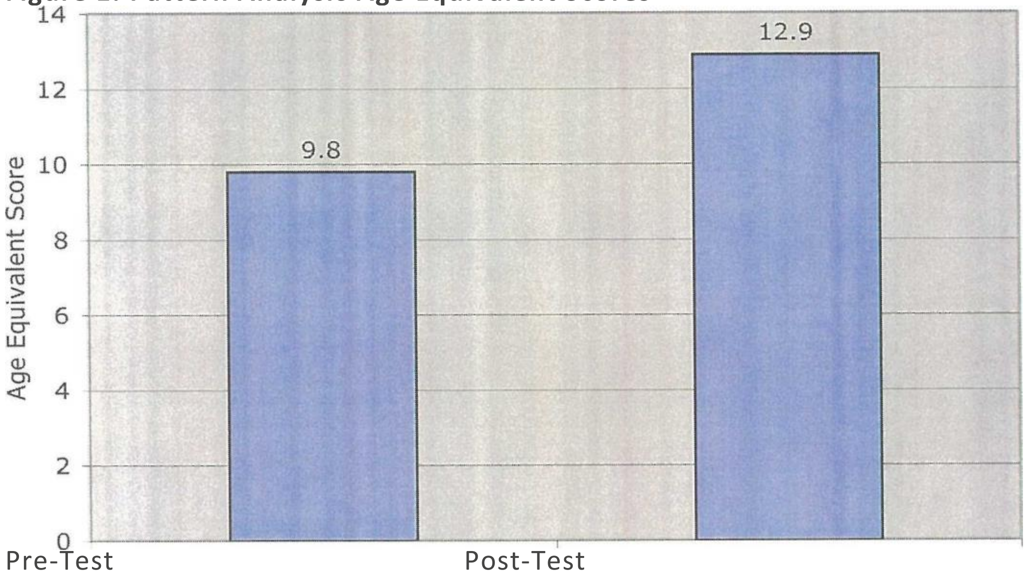
The distribution of scores on the pre- and post-tests reveals an interesting phenomenon.
None of the children achieved a raw score of forty or higher on the pre-test. Forty-two is the maximum score on this test. However, on the post-test six of the twenty-two subjects achieved a score of forty or higher, indicating that a ceiling effect came into play on this test. It is possible that if the test were harder, or had more difficult items, these subjects might have produced a higher raw score.

The Pattern Analysis test is the only one of our measures for which we can provide "age-equivalent" scores.[2] On the pre-test the sample on average scored at the 9.8 year old age level. On the post-test the sample scored at the 12.9 year old age level. This is a difference of 3.1 years.

[2] The Paper Folding test provides a table to calculate age equivalent scores. However, since some test subjects failed to answer at least one question correctly we could not compute age equivalents for this test. The other tests do not offer age equivalent scoring.

98 High Street, Motueka, 7120
info@itecnz.co.nz www.itecnz.co.nz

Figure 1: Pattern Analysis Age Equivalent Scores



We conclude that the visual-spatial intervention is associated with higher test scores on the pattern analysis test and that a ceiling effect may have prevented these measures from capturing the full degree of ability of some of these subjects after the completion of the intervention. The size of the difference scores is a bit more than three years in age equivalent ability.

Paper Folding:

Table 3 summarizes the pre- and post-test scores for the Paper Folding subtest. On average, the sample tested 2 points higher on the post-test than on the pre-test. A t-test was computed to test the null hypothesis that the post-test score was not higher than the pre-test score (at the 0.05 level). The result of the hypothesis test ($t=1.788$, $p<.05$) reached significance at the 0.05 level which means that the null hypothesis can be rejected based on this sample. We take this as evidence that the sample performed better on the post-test than on the pre-test.

Table 3: Pre-, Post-, and Delta Scores for Paper Folding

	N	Mean	SD	Median	Mill	Max	Range
Pre-Test	22	1.64	2.77	1	0	13	13
Post-Test	22	3.64	4.46	1.5	0	14	14
Delta	22	2.00	-	0.5	-	-	-

Facial Recognition:

Table 4 summarizes the pre- and post-test scores for the Facial Recognition test. On average, the sample tested 0.64 points higher on the post-test than on the pre-test. A t-test was computed to test the null hypothesis that the post-test score was not higher than the pre-test score (at the 0.05 level). The result of the hypothesis test (t—0.798, p=.21) does not reach significance at the 0.05 level which means that the null hypothesis cannot be rejected based on this sample. This result could be influenced by a ceiling effect. Only three of the subjects scored below the 85th percentile on the pre-test, not leaving much room to improve on this measure.

Table 4: Pre-, Post-, and Delta Scores for Facial Recognition

	N	Mean	SD	Median	Min	Max	Range
Pre-Test	22	18.64	2.94	19.5	11	23	12
Post-Test	22	19.27	2.31	19	15	23	8
Delta	22	0.64	-	0.5	-	-	-

Block Construction:

Table 5 summarizes the pre- and post-test scores for the Block Construction test. On average, the sample tested 2.05 points higher on the post-test than on the pre-test. A t-test was computed to test the null hypothesis that the post-test score was not higher than the pre-test score (at the 0.05 level). The result of the hypothesis test (1=1.85, p<.05) reached significance at the 0.05 level which means that the null hypothesis can be rejected based on this sample. We take this as evidence that the sample performed better on the post-test than on the pre-test. This result should be considered along with the fact that several of the subjects earned perfect scores (29) on the pre-test.

Table 5: Pre-, Post-, and Delta Scores for Block Construction

	N	Mean	SD	Median	Min	Max	Range
Pre-Test	22	24.91	4.20	27	16	29	13
Post-Test	22	26.95	3.05	28	16	29	13
Delta	22	2.05	-	1	-	—	-

Line Orientation:

Table 6 summarizes the pre- and post-test scores for the Line Orientation test. On average, the sample tested 2.55 points higher on the post-test than on the pre-test. A t-test was computed to test the null hypothesis that the post-test score was not higher than the pre-test score (at the 0.05 level). The result of the hypothesis test (t=1.15, p=0.12) failed to reach significance at the 0.05 level which means that the null hypothesis cannot be rejected based on this sample.

Table 6: Pre-, Post-, and Delta Scores for Line Orientation

	N	Mean	SD	Median	Min	Max	Range
Pre-Test	22	17.45	7.81	18.5	2	29	27
Post-Test	22	20	6.82	21.5	4	30	26
Delta	22	2.55	3.92	2,5	-6	8	14

Form. Discrimination:

Table 7 summarizes the pre- and post-test scores for the Form Discrimination test. On average, the sample tested 1.77 points higher on the post-test than on the pre-test. A t-test was computed to test the null hypothesis that the post-test score was not higher than the pre-test score (at the 0.05 level). The result of the hypothesis test (t— 1.44, p=.07) approached significance at the 0.05 level but did not meet it. Which means that the null hypothesis cannot be rejected based on this sample. This result may be a reflection of a ceiling effect. Several students received perfect scores on the pre-test.

Table 7: Pre-, Post-, and Delta Scores for Form Discrimination

	N	Mean	SD	Median	Min	Max	Range
Pre-Test	22	25.77	4.73		10	32	22
Post-Test	22	27.55	3.29	28	21	32	11
Delta	22	1.77	4.03	1.5	-9	11	20

Mental Rotations:

Table 8 summarizes the pre- and post-test scores for the Mental Rotations test. On average, the sample tested 3.64 points higher on the post-test than on the pre-test. A t-test was computed to test the null hypothesis that the post-test score was not higher than the pre-test score (at the 0.05 level). The result of the hypothesis test (t=2.33, p<.05) reached significance at the 0.05 level which means that the null hypothesis can be rejected based on this sample. We take this as evidence that the sample performed better on the post-test than on the pre-test.

Table 8: Pre-. Post-, and Delta Scores tor Mental Rotations

	N	Mean	SD	Median	Min	Max	Range
Pre-Test	22	7.36	4.499	7	0	18	18
Post-Test	22	11	5.79	10.5	2	25	23
Delta	22	3.64	5.22	4	-6	12	18

98 High Street, Motueka, 7120
info@itecnz.co.nz www.itecnz.co.nz

Main Findings

Statistically significant gains occurred in four of the seven tests in the visual spatial test-battery. Children's post test scores were significantly higher than pre-test in Pattern Analysis, Paper Folding, Block Construction and Mental Rotations.

Children in the pilot study did not show statistically significant improvements on the Facial Recognition, Line Orientation or Form Discrimination tests. However, the findings for the Facial Recognition and Form Discrimination tests are questionable as the tests did not contain enough difficult items to measure the ability level of several of the subjects. We analyzed the pre-test scores in order to determine how many subjects had room to improve on the test protocol in the post-test (See Table 8). We found that Pattern Analysis, Paper Folding and Mental Rotations seemed to be the most challenging of the tests used, and thus gave children the most room to improve on the post test. Tt is possible that the other tests were too easy to provide the children room to show improvement.

Table 9: Ceiling Effect in Test Battery

Test	Subjects scoring below 85 percent on pretest	Percentage of total sample
Pattern Analysis	21	95.5%
Paper Folding	21	95.5%
Facial Recognition	3	13.6%
Block Construction	5	22.7%
Line Orientation	16	72.7%
Form Discrimination	9	40.9%
Mental Rotations	22	100%

Other Findings

Post-test scores on the visual-spatial subtests of the Stanford Binet (Pattern Analysis and Paper Folding) were strongly correlated with the achievement level of the children using the software. We found statistically significant correlations between the variable representing the highest level achieved with each of the software activities with the post-test scores on the two Stanford-Binet subtests. We understand this to be evidence that the software activities address skills related to these two test measures of visual-spatial reasoning.

Children with learning disabilities as well ADD or ADHD (parent report) did not show as much improvement on the test battery as children with only learning disabilities. This might have to do with the fact that children who struggle with attention issues were not able to sustain their focus on the software activities during the long computer sessions (2-hours each) and would be better suited to using the software for shorter time periods.

Strengths and Weaknesses of the Pilot Study

There are considerable weaknesses to this pilot study that limit the strength of the key findings. The main weakness is that there was not a control group who also took the pre-and post-tests. Such a control would provide a baseline to which we could compare the gains made by the children who used the software over the course of the summer. Some of the gains made by the subjects in this pilot study could be attributed to familiarity with the test and the examiner at the time of the post-test.

In addition, using only tests of visual spatial skills limited our ability to see any transfer of skills to other realms. Subsequent testing will use a full complement of neuropsychological tests in order to determine if use of the software could be associated with any changes in other abilities such as logical reasoning or vocabulary. Some of the measures that we used were of limited value because they were not challenging enough for some children in the pilot study. This resulted in a ceiling effect that prevented us from capturing the highest level of performance for some subjects. The sample was small and was not a diverse group of children. This group of children was representative of a small segment of the population, specifically middle and upper class, suburban, Caucasian families. The majority of the sample was boys

The study staff spent a lot of time assisting the children, explaining the software, giving instructions and helping fix bugs. The software was new when this study was conducted and therefore had usability issues that needed attention. As the software is refined and improved, there will be less need for staff intervention and we will be able to investigate whether or not this intervention can be staffed by individuals who do not specialize in Cognitive Science.

Conclusion

This pilot study represents an early test for the development of a new technology to develop cognitive abilities. The data presented her lends support to the concept that software can be designed to efficiently deliver cognitive interventions to students. The data is encouraging enough to offer cause to continue with the development of the technology and to pursue more rigorous evaluation of its effectiveness. These results are exciting but we cannot lose sight of the limitations of the pilot study - the primary one being that there is no control group by which to compare the gains on the post-test measures. Subsequent outcome tests for this project will include a control group of evenly matched children that will provide a baseline from which to compare any potential improvements in student's cognitive abilities as well as evaluations in schools that consider possible effects on school learning.

The primary question driving this research is whether technology can be used to make cognitive interventions more widely available to our nations school children. The pilot study summarized in this paper provides early evidence to support the argument that technology can be harnessed to overcome the barriers to scalability and replication that have plagued these kinds of specialized interventions. Future testing will involve more mature evolutions of the software as well as more rigorous experimental methodologies.

References

Adey, P., & Shayer, M. (1993). An exploration of long-term far-transfer effects following an extended intervention program in the high school science curriculum. *Cognition and Instruction, 11(1),* 1-29.

Adey, P., & Shayer, M. (1994). *Really raising standards: Cognitive intervention and academic achievement.* New York: Routledge.

Adey, P. S. (1988). Cognitive acceleration - review and prospects. *International Journal of Science Education, 10(2),* 121-134.

Blagg, N. (1991). *Can we teach intelligence? A comprehensive evaluation of Feuerstein's Instrumental Enrichment program.* Hillsdale, NJ: Eribaum.

Bond, T.G. and C.M. Fox, Applying the rasch model: Fundamental measurement in the human sciences. 2001, Mahwah, NJ: Lawrence Eribaum.

Brown, A. (1989). Analogical learning and transfer: What develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical reasoning* (pp. 369-412). Cambridge, MA: Cambridge University Press.

Brown, A. L. (1990). Domain-specific principles affect learning and transfer. *Cognitive Science, 14,* 107-133.

Case, R. (1998). The Development of Conceptual Structures. In D. Kuhn & R. S. Siegler (Eds.), *Cognition, Perception, and Language.* (Fifth ed., Vol. 2, pp. 745-800). New York: John Wiley & Sons, Inc.

Case, R., & Griffin, S. (1990). Child cognitive development: The role of central conceptual structure in the development of scientific and social thought. In C. A. Haver (Ed.), *Advances in psychology: Developmental psychology* (pp. 193-230). North Holland: Elsevier.

Case, R., & McKeough, A. (1990). Schooling and the development of central conceptual structures: An example from the domain of children's narrative. *International Journal of Educational Psychology, 8,* 835-855.

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development.* 60(5/6).

Connell, M. W. (2002). *On the viability of computational neuroscience as a framework for connection brain, mind, and education.* Unpublished manuscript. Available from Harvard Graduate School of Education, Cambridge, MA 02138.

Connell, M. W. & Stevens, D. A., "A Computer-Based Tutoring System for Visual-Spatial Skills: Dynamically Adapting to the User's Developmental Range", Proceedings of the 2nd International Conference on Development and Learning (IC.DL 2002), IEEE Computer Society Press, Los Alamitos, Calif., 2002, pp. 245-251.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920.* New York: Teachers College Press.

Cuban, L. (2001). *Oversold and underused: Computers in classrooms.* Cambridge, MA: Harvard University Press.

Dawson, T. L. (2001, June). *Meaning language and stage:* A *new look at moral thought.* Paper presented at the Annual Meeting of the Jean Piaget Society, Berkeley, CA.

Detterman, D. K. (1993). The case for the prosecution: Transfer as an Epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on Trial: Intelligence, Cognition, and Instruction .* Norwood, N.T: Ablex.

Eastman, P. M. and Carry, L. R. (1975). Interaction of spatial visualization and general

reasoning abilities with instructional treatment in quadratic inequalities: A further investigation. *Journal for Research in Mathematics Education (6)3,* pp. 142-149.
Eliot, J. and Hauptman, A. (1981). Different dimensions of spatial ability. *Studies in Science Education, 8,* pp. 45-66.
Feuerstein. R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability.* Baltimore: University Park Press.
Feuerstein, R., Miller, R., Hoffman, M. B., Rand, Y., Mintzker, Y., & Jensen, M. (1981). Cognitive modifiability in adolescence: Cognitive structure and the effects of intervention. *The Journal of Special Education, 15,* 269-286.
Feuerstein, R. (1980). *Instructional enrichment: An intervention program for cognitive modifiability.* Baltimore, MD: Park Press.
Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87,* 477-531.
Fischer, K. W., & Farrar, M. J. (1987). Generalizations about Generalization: How a theory of skill development explains both generality and specificity. *International Journal of Psychology, 22,* 643-677.
Gagnon, D. (1985). Videogames and spatial skills: An exploratory study. *Educational Communication and Technology Journal 33(4),* pp. 263-275.
Griffin. S., & Case, R. (1996). Evaluating the Breadth and Depth of Training Effects When Central Conceptual Structures Are Taught. *Monographs of the Society for Research in Child Development, 60(5/6),* 83-101.
Griffin, S. A., Case, R., & Capodilupo, S. (1995). Teaching for understanding: The importance of central conceptual structures in the elementary school mathematics curriculum. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Postering generalization in learning. .* Hillsdale, NJ: Erlbaum.
Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice .* Cambridge: MIT Press.
Guilford, J.P. (1967). *The Nature of Human Intelligence..* New York: McGraw-Hill.
Haywood, M. and Coltheart, M. (2000). Neglect dyslexia and the early stages of visual word recognition. *Neurocase 6(1),* pp. 33-44.
Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press.
Hill, D. M. and Oberhauf, P. A. (1979). Spatial visualization, problem solving, and cognitive development in freshman teacher education students. *Science. Education a63(5),* pp. 665-670.
Kalchman, M. (2001). *Teaching Mathematical Functions: Applying Case's Developmental Theory to Classroom Instruction.* Paper presented at the AERA annual meeting, Seattle, WA. *Kassubek, J., Schmidtke, K., Kimming, H., Lucking, C. H., Greenlee, M. W. (2001).*
Changes in cortical activation during mirror reading before and after training: an fMRI study of procedural learning. Cognitive Brain Research, 10, *pp. 207-217.*
Lowery, B. R. and Knirlc, F. G. Micro-computer video games and spatial visualization acquisition. *Journal of Educational Technology Systems 11(2),* pp. 155-166.
McClurg, P. A. (1992). Investigating the development of spatial cognition in problem-solving microworlds. *Journal of Computing in Childhood Education, 3.* pp. 111-126.
McCormack. A. (1988). *Visual/spatial thinking: An essential element of elementary school science.* Monograph and Occasional Paper Series, Council for Elementary Science International.

McGilly, Kate (Ed.). (1994). *Classroom lessons: Integrating cognitive theory and classroom practice.* Cambridge, MA: MIT Press.
Meeker, M.N. (1969). *The structure of intellect.* Columbus, OH: Merrill.
Moses, B. (1979). *The effects of spatial instruction on mathematical problem solving performance.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.
Neisser, U. (Ed.). (1998). *The rising curve: long-term gains in IQ and related, measures.* Washington, DC: American Psychological Association.
Perkins, D. N. (1995). *Outsmarting IQ: the emerging science of learnable intelligence.* New York: Free Press.
Perkins, D. N., & Grotzer, T. S. (1997). Teaching Intelligence. *American Psychologist.* 52(10), 1125-1133.
Perkins, D. N., & Salomon, G. (1989). Are Cognitive Skills Context-Bound? *Educational Researcher, 18(* 1), 16-25.
Piaget, J. (1983). Piagefs Theory. In P. Mussen (Ed.), *Handbook of Child Psychology: Vol. 1* (Vol. 1,). New York: Wiley.
Piaget, J. (1975/1985). *The Equilibration of Cognitive Structures: The central problem of intellectual development.* Chicago: University of Chicago Press.
Pillsbury. (1908). The effects of training on memory. *Educational Review. 36,* 15-27.
Rand, Y., Tannenbaum, A. J., & Feuerstein, R. (1979). Effects of Instrumental Enrichment on the psycho educational development of low-functioning adolescents. *Journal of Educational Psychology, 71,* 751-763.
Salomon, G., & Perkins, D. N. (1987). Transfer of cognitive skills from programming: When and how? *Journal of Educational Computing Research, 3,* 149-169.
Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24,* 185-221.
Shayer, M. (1988). Neo-Piagetian theories and educational practice. In A. Demetriou (Ed.), *The Neo-Piagetian theories of cognitive development: Towards cm Integration.* . Amsterdam: North Holland.
Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill.* Cambridge, MA: Harvard University Press.
Stankov, L. (1986). Kvashchev's Experiment: Can We Boost Intelligence? *Intelligence. 10,* 209-230.
Stankov, L. (1991). The Effects of Training and Practice on Human Abilities .
Stankov, L., & Chen, K. (1988). Training and Changes in Fluid and Crystallized Intelligence. *Contemporary Educational Psychology, 13,* 382-397. Sternberg, R. J. How can we teach intelligence? *Educational Leadership,* 42(1), 38-50.
Sternberg, R. J. (1988). The Raising of Intelligence: A Selected history of Attempts to Raise Retarded Intelligence. *American Journal of Psychology, 101(1),* 142-147.
Sternberg, R. J. (1992). *Can intelligence be increased.* Guilford, CT: Dushkin Publishing Group.
Sternberg, R. J.. & Bhana, K. (1986). Synthesis of Research on the Effectiveness of Intellectual Skills Programs: Snake-oil Remedies or Miracle Cures? *Educational Leadership,* 44(2), 60-67.
Sternberg, R. J., & Detterman, D. K. (1982). *How and how much can intelligence be increased?* Norwood, NJ: Ablex.
Sternberg, R. J., & French, P. A. (1993). Mechanisms of Transfer. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on Trial: Intelligence, cognition, and instruction .* Norwood, NJ: Ablex.
Stevens, D. A. (1999). *Transfer revisited: A critical review.* Unpublished manuscript.

Available from Harvard Graduate School of Education, Cambridge, MA 02138.

Stevens, D. A. (2000). *Piaget as a dynamic systems theorist.* , Harvard University Doctoral Dissertation, Cambridge, MA.

Stevens. D. A. (2000). *Dynamic systems; Continuities and discontinuities with Piaget's theory of equilibration.* Paper presented at the Annual Meeting of the Jean Piaget Society, Montreal, Canada.

Stevens, D. A., Connell, W. M., Pardi, R.. Pilgrim, B., Schwarz, P. (2002). *Designing Educational Software to Improve Visual-Spatial Ability: Implications for Theory, Measurement, and Design.* Paper presented at the Annual Meeting of the Jean Piaget Society, Philadelphia, PA.

Subrahmanyam, K. and Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology 15,* pp. 13-32.

Tallal, P. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271.*

Theisen, Patricia J. (1998). Perceptual Enrichment Program (PEP): Resource Handbook. Seattle, WA.

Wachs, H. (2000). Visual-Spatial Thinking. The Interdisciplinary Council on Developmental and Teaming Disorders; Clinical Practice Guidelines. Bethesda, MD: 1CDL Press.

98 High Street, Motueka, 7120
info@itecnz.co.nz www.itecnz.co.nz