COGNITIVE EFFECTS OF CHESS INSTRUCTION ON STUDENTS AT RISK FOR ACADEMIC FAILURE

Saahoon Hong
and
William M. Bart

University Of Minnesota

Cognitive effects of chess instruction on students at risk for academic failure was examined. Thirty-eight students, from three elementary schools, participated in this study. The experimental group received a ninety-minute chess lesson once per week over a three-month period; and the control group students regularly attended school activities after class. The experimental group performance on the test was not different from the control group performance. However, chess skill rating and TONI-3 posttest scores were significantly correlated when controlling for TONI-3 pretest score (d = 0.29). This suggests that chess skill rating is a key predictor for the improvement of student cognitive skills. Students at risk at beginning levels of competency in chess may be able to improve their cognitive skills and to improve their skill at chess.

Chess playing is a strategy game that requires higher order cognitive skills. The acquisition of higher order cognitive skills plays a major role in enabling students to better establish and attain goals, identify potential responses when making decisions, and achieve self-regulated learning (Wehmeyer, Palmer, Agran, Mithaug, & Martin, 2000). As a result, investigators have examined the usefulness of chess playing to develop higher order cognitive skills (Horgan, 1987; Horgan & Morgan, 1990). Higher order cognitive skills such as analysis, evaluation, and logical thinking are prevalent in the game of chess (Grossen, 1991).

Chess playing involves the comprehension of chess positions, the analysis of moves and their sequences, and the evaluation of positions resulting from certain moves (Bart, 2004; Cleveland, 1907; Gobet & Simon, 1996; Holding, 1985). Since these processes are considered to be transferable skills (Ericsson & Staszewski, 1989; Ericsson & Kintsch, 1995; Gobet & Simon, 1996), chess playing receives considerable attention as a learning tool and part of the curriculum.

Research on chess instruction has tended to provide empirical support for the beneficial effects of chess training on performance on cognitive tasks (Horgan, 1987; Smith & Cage, 2000; Christiaen & Verholfstadt 1978; Frank & D’Hondt, 1979). For example, in an experimental study, Frank & D’Hordt (1979) found that an experimental group of learners receiving chess instruction scored better on both numerical and verbal aptitude tests than did a control group of learners not receiving chess instruction. These findings lend credence to the application of chess instruction to students with cognitive challenges. Thus chess instruction may be a productive intervention for students at risk for academic failure.

Students at risk are defined as students who are one or more years behind their age or grade level in mathematics or reading skills (Sapp, 1993). Most of them require the same assistance as students with disabilities (Sapp & Farrell, 1994). Students at risk tend to rely on previously employed but unsuccessful responses, process information less effectively, and are often unable to solve problems in their lives (Agran & Wehmeyer, 1999; Swanson & Alexander, 1997; Wehmeyer & Keleher, 1995). They have difficulties in utilizing higher order cognitive skills. Feuerstein (1980) claimed that enriched environments could resolve these difficulties. Teaching and practicing these cognitive skills through chess playing to students at risk have produced better results in basic skills than over reliance on drilling, direct instruction, or other current school improvement methodologies (Pogrow, 1988). Pogrow even argued that the acquisition of higher order cognitive
skills compensates students at risk who are deficient in basic skills, because higher order cognitive skills are considered as a knowledge base for all learning. In addition, Pogrow contended that students at risk have competencies to provide solutions on even difficult tasks requiring higher order thinking processes, when enough time and resources are given.

However, this hardly occurs in education for students at risk, because they are provided with less opportunity to improve higher order cognitive skills (Allington & McGill-Branzen, 1989). This lack of instructional opportunities resulted from the view that students at risk could not benefit from instruction in higher order cognitive skills (Leshowitz, Jenkens, Heaton, & Bough, 1993).

A common approach for students at risk is to remedy their deficiencies in the basics, like reading, writing, and math. This approach mostly relies on repetitive drill. Knapp and Shields (1990) criticized the repetitive drill approach that tends to: (a) underestimate student competencies; (b) prevent students from accessing more challenging and interesting work; and (c) deprive students of a meaningful context for learning. Such criticism sheds light on the development of higher order cognitive skill instruction (Means & Knapp, 1991). Pogrow’s model supports the view that teaching higher order cognitive skills provides students at risk with opportunities to use what they already know, in the form of encoding and retrieving processes. Consequently, these processes could lead students at risk to major gains in basic skills.

In conclusion, research on chess instruction for students at risk may likely provide both regular and special educators with practical suggestions on how to develop higher order cognitive skills and to improve scholastic achievement levels among learners. Furthermore, Storey (2000) suggested that chess instruction could also benefit children with disabilities, even though only anecdotal evidence is available for the effect of chess play on students with disabilities (Remsen, 1998; Wojcio, 1995). This study will examine this issue as it concerns students who are at risk for academic failure. The main purpose of this study is to examine cognitive effects of chess instruction on students at risk for academic failure.

**Method**

**Participants**

Thirty-eight students, ages 8 to 12, from three elementary schools participated in this study. The schools are located in Seoul, Korea. There were 20 students from one school, seven students from a second school, and eleven students from a third school.

The students were randomly placed into two groups: a control group and an experimental group. There were 15 males and 5 females in the control group with an average age of 9.74 years and 12 males and 6 females in the experimental group with an average age of 9.71 years. In the control group, there were 17 students at risk and 3 students with learning disabilities and, in the experimental group, there were 15 students at risk and 3 students with learning disabilities. As to the distribution of students by grade, the control group consisted of three students in third grade, nine students in fourth grade, seven students in fifth grade, and one student in sixth grade. The experimental group consisted of three students were in third grade, five students in fourth grade, six students in fifth grade, and four students in sixth grade.

**Instruments**

The Korean Basic Skills Test. The Korean Ministry of Education and Human Resource Development and the Korean Institute of Curriculum and Evaluation in 2002 collaboratively developed the Korean Basic Skills Test (KBST) in 2002. The KBST measures student basic abilities in mathematics, reading, and writing. For third grade students (Ministry of Education and Human Resource Development, 2003), the average KBST scores were 93.89 for reading, 94.88 for writing, and 92.28 for mathematics. The basic ability cutoff scores for students at risk were 75 for reading, 78 for writing, and 77 for math. The percentages of students below those cutoff scores were 3.45 percent for reading, 3.00 percent for writing, and 6.84 percent for mathematics. 1.34 percent of students were identified as student at risk in all reading, writing, and mathematics. Student at risk for academic failure lacked basic abilities in reading, mathematics, or writing.

One investigator identified students at risk by using the KBST. Approximately 3-5% of the students per school fell into this category. The students at risk showed significant deficits in more than one area among the domains of reading, writing, and mathematics.
The Raven’s Progressive Matrices Test. The Raven’s Progressive Matrices Test (RPM) is designed to measure nonverbal abilities such as student perception of relationships in geometric figures and reasoning by analogy independent of language and formal schooling (Raven, Raven, & Court, 2000). The RPM is also considered to be a fine measure of logical ability and spatial ability (Raven, Court, and Raven, 1985). The RPM comes in three types: the Colored Progressive Matrices (CPM), the Standard Progressive Matrices (SPM), and the Advanced Progressive Matrices (APM). This study made use of the SPM consisting of five sets with 12 problems in each set. The test-retest reliabilities range from .83 to .93. In this study, the correlation between the RPM pretest scores and the posttest scores was .78.

The SPM has fine concurrent validity. For example, correlations between the SPM and WISC-R ranged from .83 to .92 in a stratified sample of Canadian children ranging in age from seven to eleven years (Rogers & Holmes, 1987). Horgan and Morgan (1990) contended that the type of reasoning required to solve SPM items is similar to chess reasoning for a wide range of subjects, even though norms for children are limited.

The Test of Nonverbal Intelligence – Third Edition. The Test of Nonverbal Intelligence-Third Edition (TONI-3) is a norm-referenced test and a language-free measure of cognitive ability (Brown, Sherbenou, & Johnsoen, 1997). In particular, the TONI-3 was designed to measure problem solving, aptitude, and reasoning skills. Two equivalent forms are available. Each form of the TONI-3 has 50 items. Converted scores from obtained raw scores are provided with a mean 100 and a standard deviation of 15. It is particularly useful for individuals who are believed to have difficulties in taking tests, disabilities, or lack of exposure to the British and United States cultures. In this study, students received two forms (A and B).

The TONI-3 has fine psychometric properties. For example, alternate forms reliability has ranged from .79 to .92. A correlation between TONI and SPM was .92. In this study, the correlation between the two forms of the TONI-3 tests was .69.

Chess Quiz. Students in the experimental group received a Chess Quiz that was developed by the chess instructor. At the twentieth session of chess instruction, students in the experimental group completed the Chess Quiz with a score range of 0 to 40. For each item, each participant in the experimental group was asked to find all possible capture moves in a position. The Chess Quiz did not include any questions assessing the use of chess strategies. It simply measured student knowledge about chess pieces and their moves. The Pearson product-moment correlation between chess class attendance and the quiz was .80.

Chess Skill Rating. Chessmaster 9000 provided artificial opponents to participants in the experimental group. With each game of chess, a participant received a score of 1 for a win, a score of .5 for a draw or a stalemate, and a score of 0 for a loss. Each artificial player had an Elo rating that indicated the quality of its play. Elo ratings range from 0 for a beginner to approximately 2850 for a World Chess Champion. The Elo rating scale is the official scale of chess player skill for the United States Chess Federation.

On the basis of the Elo scale, players with ratings of 2500 and above are called grandmasters, and players with ratings of below 1200, Class E. Until the players complete 20 games, they are given a provisional rating. Each participant in the experimental group played against an initial artificial opponent provided by Chessmaster 9000 that had a chess rating of 300. The formula for chess skill rating presented by Fogel, Hays, Hahn, and Quon (2004) was used to determine the Elo rating of each participant.

Procedure
After each school identified the students at risk, the homeroom teacher sent consent forms to the students and their parents. The study began with administration of two pretests after the consent forms were returned. A researcher and a research assistant administered the TONI-3 and the RPM to the students in the first week of this study. The TONI-3 was administered individually and the RPM was administered in groups.

The participants were then randomly assigned to an experimental group or a control group. The experimental group received a 90-minute chess lesson once per week and the control group students
attended regular school activities after class. At the end of the chess intervention, the participants received the TONI-3 and the RPM. Students in the experimental group completed the Chess Quiz.

Chess instruction consisted of 12 separate lessons over a 3-month period. Each lesson included three segments: reviewing, lecturing, and chess playing. The chess instructor developed and provided a set of quizzes. The quiz was used to identify student difficulty in understanding chess moves and rules. Each subsequent lesson started with reviewing a previous lesson and a quiz. The last six lessons were implemented in a computer lab with chess software and allowed students to practice higher order cognitive skills.

Overall, the chess instructor asked the students to follow four steps to develop their chess skills: (1) understand chess rules; (2) think ahead for a plan; (3) implement the plan; and (4) seek feedback and rehearsal. The researcher and the chess instructor developed twelve sessions derived from the Comprehensive Chess Course (Pelts & Alburt, 1992). The chess software was used as a tool to practice and generalize the contents of each lesson.

Chess playing was new to most of the students. Although three students stated that they sometimes played chess with their brothers, their knowledge of basic chess rules was shallow.

Results

Preliminary Analysis

Although students were assigned randomly to each group, the pre-test intelligence test means for the control and experimental groups were compared using a one-way analysis of variance (ANOVA). The groups appeared equivalent on gender, age, grade, school, and disabilities. The mean TONI-3 pretest scores of the control group (M = 85.60; SD = 20.48) and the experimental group (M = 96.50; SD = 17.12) were not significantly different, F (1, 36) = 3.13, p > .05. The mean Raven’s Progressive Matrices (RPM) pretest scores of the control group (M = 26.20; SD = 10.96) and the experimental group (M = 29.39; SD = 8.56) were also not significantly different, F (1, 36) = .98, p > .05.

In addition, comparable KBST pretest mean scores for the two groups were not significantly different. For the reading KBST pretest scores, the means of the control group (M = 59.08; SD = 21.72, n = 12) and the experimental group (M = 66.71; SD = 15.68, n = 7) were not significantly different with F (1, 17) = .66, p > .05. For the mathematics KBST pretest scores, the means of the control group (M = 60.00; SD = 16.78, n = 11) and the experimental group (M = 64.29; SD = 9.27, n = 9) were not significantly different with F (1, 16) = .38, p > .05. For the writing KBST pretest scores, the means of the control group (M = 68.56; SD = 17.66, n = 9) and the experimental group (M = 62.20; SD = 21.95, n = 5) were also not significantly different with F (1, 12) = .35, p > .05.

Some Descriptive and Inferential Statistical Analyses

Table 1 provides descriptive statistics for the intelligence test scores.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Group</th>
<th>Pre-Test M</th>
<th>Pre-Test SD</th>
<th>Post-Test M</th>
<th>Post-Test SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TONI-3</td>
<td>Control</td>
<td>85.60</td>
<td>20.49</td>
<td>97.25</td>
<td>13.18</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>96.50</td>
<td>17.12</td>
<td>100.83</td>
<td>11.78</td>
</tr>
<tr>
<td>RPM</td>
<td>Control</td>
<td>32.30</td>
<td>27.60</td>
<td>39.20</td>
<td>28.66</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>37.33</td>
<td>26.03</td>
<td>40.94</td>
<td>23.31</td>
</tr>
</tbody>
</table>

A repeated measures ANOVA with a 2 X 2 factorial design was employed to determine whether chess instruction would influence the experimental group TONI-3 and RPM scores. The first factor related to treatment, i.e., the control and experimental groups. The second factor represented time, i.e., the pretest and the posttest. The repeated measures ANOVA examine the main effect and the interactive effect of treatment and time as independent variables on the TONI-3 and RPM scores as dependent variables.

Although the TONI-3 mean scores significantly increased from the pretest to the posttest, F (1, 36) = 11.84, p < .001, the main effect for chess instruction was not significant for the TONI-3 with F (1, 36) = 2.40, p > .05. The treatment X time interaction effect, reflecting differences among the groups in amount of change, was also not statistically significant for the TONI-3 with F (1, 36) = 2.481, p > .05. In other words, the changed scores of TONI-3 in the experimental chess group were similar to those in
the control group. There was no significant difference between the two groups after the chess instruction. The effect size for the experimental group was 0.29 and the effect size for the control group was 0.68.

As to the results of the repeated measures ANOVA on the other intelligence test scores, the RPM mean scores significantly increased from the pretest to the posttest with $F(1, 36) = 4.20, p < .05$. But the results of repeated measures of ANOVA show that the main effect for treatment was not significant for the RPM with $F (1, 36) = .169, p > .05$ and the treatment X time interaction effect, reflecting differences among the groups in amount of change, was also not significant for the RPM with $F (1, 33) = .756, p > .05$. In other words, the score changes for the RPM in the chess group were similar to those of the control group. There was no significant difference between the two groups after the chess instruction. The effect size for the experimental group was 0.15 and the effect size for the control group was 0.25.

Table 2 provides descriptive statistics for chess-related measures.

<table>
<thead>
<tr>
<th>Measure</th>
<th>M</th>
<th>SD</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chess Skill Rating</td>
<td>131.39</td>
<td>84.94</td>
<td>441</td>
<td>101</td>
</tr>
<tr>
<td>Chess Quiz Score</td>
<td>22.83</td>
<td>11.29</td>
<td>39</td>
<td>4</td>
</tr>
<tr>
<td>Chess Practice in Minutes</td>
<td>620</td>
<td>194.97</td>
<td>900</td>
<td>270</td>
</tr>
</tbody>
</table>

Note. Highest Possible Chess Quiz Score = 40. Highest Possible Chess Practice Score = 1080 minutes.

Regarding the chess quiz with its score range of 0-40, some students scored very well on the chess quiz and other students scored rather poorly. The maximum chess quiz score that a student received was 39; whereas, the minimum was 4. The maximum minutes of chess practice was 900 minutes; whereas, the minimum was 270 minutes. Student practice length outside chess class was not counted in this study.

Partial Correlation Analysis

Partial correlation analysis was used to explore relationships among pre- and posttest scores, chess skill ratings, chess quiz scores, and chess practice for participants in the experimental group. Table 3 provides the partial correlations among those variables controlling for TONI-3 pretest scores.

<table>
<thead>
<tr>
<th>Control Variable</th>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>TONI-3 Pretest</td>
<td>1. TONI-3 Posttest</td>
<td>1.00</td>
<td>.52*</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>2. Chess Skill Rating</td>
<td>1.00</td>
<td>.42</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>3. Chess Quiz</td>
<td>1.00</td>
<td>.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Chess Practice</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

*p < .05

Among those partial correlations, only the partial correlation between the TONI-3 posttest score and chess skill rating controlling for TONI-3 pretest score was significant with $r = .52, p < .05$.

The median of TONI-3 pretest scores divided the TONI-3 pretest scores into a Low group and a High group. All students showing improvement in chess skill ratings were in the High group of TONI-3 pretest scores. Thus, student TONI-3 posttest scores in the High group are somewhat related to chess skill ratings. In the Low group of TONI-3 pretest, chess skill ratings remained the same.

A stepwise regression was conducted to evaluate whether variables, like TONI-3 pretest score, chess skill rating, chess quiz, and chess practice, were necessary to predict TONI-3 posttest score. The stepwise regression analysis indicated a model that included two significant predictors, $F (2, 15) = 12.25, p < .001$. The two predictors, TONI-3 pretest score and chess skill rating, were positively associated with the TONI-3 posttest scores. The TONI-3 pretest score is a predictor for the TONI-3 post-test score ($R^2 = .480$). $R^2$ changes to .620 with the addition of the chess skill rating. They account for 62% of the variance among the TONI-3 posttest scores. Although the sample size is small, this result suggests that the chess skill rating was somewhat related to the increased posttest TONI-3 scores.
In contrast, the partial correlation of the RPM posttest score and the chess skill rating with the RPM pretest score being held constant was not significant, \( r = .11 \). Table 4 provides the partial correlations among selected variables controlling for RPM pretest scores.

**Table 4**

<table>
<thead>
<tr>
<th>Control Variable</th>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPM Pretest</td>
<td>1. RPM Posttest</td>
<td>1.00</td>
<td>.11</td>
<td>.03</td>
<td>.17</td>
</tr>
<tr>
<td></td>
<td>2. Chess Skill Rating</td>
<td>1.00</td>
<td>.31</td>
<td>.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Chess Quiz</td>
<td></td>
<td>1.00</td>
<td>.50*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Chess Practice</td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

\(*) p < .05\)

**Discussion**

The results of this study indicate a lack of cognitive effects of chess instruction. In the analysis of two cognitive tests, changes in experimental group performances were not different from changes in the control group performances. The results tend not to support the view that chess instruction for the beginner at risk for academic failure has salutary cognitive effects on such students. This finding is not consistent with the results of previous studies (Christiaen & Verholfstadt, 1978; Frank & D’Hondt, 1979; Smith & Cage, 2000) that showed improved cognitive skills after providing chess instruction.

This inconsistent result could be explained by two interpretations:

The first interpretation is that students at risk could require more time for chess instruction than a twelve-session chess instruction period for one semester. Pogrow (1988) held that time and resources are key factors in developing higher reasoning skills. Students at risk could require more sessions to develop their chess skill. Thus, the lack of cognitive effects of chess instruction might be explained by the limited number of chess instruction periods. Bart (2004) suggested at least one whole academic year and preferably two academic years as the duration for effective chess instruction. It is likely that more time on task learning chess and studying chess could facilitate the development of cognitive skills and capabilities among learners including students at risk.

The second interpretation is that novice chess players at risk for academic failure could hardly develop their cognitive skills until they reach a certain level of chess skill. This interpretation is consistent with the results of Horgan and Morgan’s (1990) study. To Horgan and Morgan, attaining certain levels of chess skill could be associated with improvement in higher order cognitive skills.

There was no correlation between chess skill rating and RPM score. That finding is not consistent with the findings of Horgan and Morgan (1990) and Frydman and Lynn (1992). This inconsistent result can be explained by different chess skill levels. In this study, the mean chess skill rating was 131.39 with a standard deviation of 84.94. The Horgan and Morgan study and the Frydman and Lynn study were conducted with club chess players with chess ratings greater than 1000.

One intriguing result in this study is that chess skill rating and TONI-3 posttest score were significantly correlated when controlling for TONI-3 pretest score. This suggests that chess skill rating is a key predictor for the improvement of student cognitive skills. Students at risk who are at beginning levels of competency in chess may be able to improve their cognitive skills and their skill at chess.

One limitation of this study is that the chess instruction suggested by Pelts and Alburt (1992) was not specifically developed for students at risk or with disabilities whose needs are individually different. A preferred model of chess instruction may focus on more opportunities for the students to acquire knowledge of strategies and tactics in chess. It is likely that the deeper levels of chess competency involving knowledge of strategies and tactics need to be acquired in order for higher levels of nonverbal intelligence and other cognitive capabilities to be attained. It is obvious that chess instruction should take the characteristics of students at risk into consideration and be reorganized for further studies. In addition, instructor knowledge of pedagogy for students at risk may contribute to effective chess instruction.

In conclusion, we recommend that the cognitive effects of chess instruction on students at risk for academic failure continue to be studied. Chess instruction specially configured may prove to be very
efficacious in producing salutary cognitive effects among students at risk for academic failure in the USA, and elsewhere in the world.

References
Cleveland, A. (1907). The psychology of chess and of learning to play it. American Journal of Psychology, 18, 269-308.


