EXPERIMENTS IN CHESS COGNITION

Danny Kopec
Monroe Newborn
Winston Yu

March , 1984

ABSTRACT

The results of three related studies measuring the performance of humans and computer chess programs on different test positions are analyzed. The experimental work comprising these studies is:

1. The Pairs Experiment
2. Computer Results
3. The Time Sequence Experiment

The underlying theme of all the work reported is oriented towards addressing the question "How does the performance of humans and computer chess programs on a set of problem positions vary with time?" There are some surprisingly tangible results.

INTRODUCTION

OSTRICH

In general there are two methods of remedy when a computer can no longer handle the complexity of a given problem. The more common, simpler solution is to switch to a more powerful machine. However this approach is not always feasible, and the fastest computer may not be fast enough for a given task.

The second solution is to employ more than one processor of the same computer family to work on a given task. In the ideal situation, the number of processors employed is inversely proportional to the required computation time. However, the processors must communicate with each other and the sub-processes must be synchronized. Thus as the number of processors increases, the margin of their increase in computation power decreases. The program OSTRICH (Newborn, 1982) uses eight Data General Nova computers in parallel. Each Nova computer is able to search one lookahead subtree at the same time. The master processor receives results from all the other processors and then selects the best one.

45

The Bratko - Kopec Experiment

The status of today's top computer chess programs, at just below the master level is primarily a result of their ability to efficiently search the lookahead tree to compute their moves (Kopec and Bratko, 1982). On the other hand their chess specific knowledge is very limited. The experimental work reported here is based upon a discrimination of two fundamentally different classes of moves in chess:

(1) tactical moves in which a lack of chess knowledge may be compensated for by additional computation.
(2) positional moves where a lack of chess knowledge cannot be compensated for.

Tactical moves include:

(a) checkmate or gain of material and/or
(b) a distinct improvement in terms of positional ends (e.g. mobility) and/or
(c) the defense to specific threats in term of (a) and (b) above.

One type of positional move is called a 'lever'. A lever is a pawn move which offers (a) an exchange with another pawn, (b) leads to an ultimate improvement in the pawn structure of the side playing it and/or (c) damages the opponent's pawn structure.

In the Bratko - Kopec experiment (ibid.), 35 human chess-player subjects and 17 computer chess programs were tested on 24 chess positions (12 tactical (T), and 12 lever (L) positions). The experiment provided quantitative evidence fully supporting the hypothesis that computer programs perform better in type T positions than in type L positions. Computer programs also performed better in T positions than humans of the same rating.

THE PAIRS EXPERIMENT

The Experimental Objective

This experiment is an extension of Bratko-Kopec experiment, though here only human chessplayer subjects were tested. The object of the experiment was to determine whether two human subjects (of approximately the same rating) can perform better than one human subject on a set of test positions whose solutions are exclusively T and L moves.

The Experimental Design

There are 58 positions (29 T, and 29 L) in this experiment, (sources are listed in the Appendix A). Twenty-two of the 58 positions were taken from the Bratko-Kopec experiment (see Kopec and Bratko, 1982, for the actual positions, their discussion and solutions) with three new positions added to comprise the test set for Part 2 of this experiment. Twenty-four new positions from a number of sources plus one from the original Bratko-Kopec test comprise Part 3 of the experiment.

The experiment is divided into 3 parts. In Part 1 each subject is given 8 practice positions. In Part 2 and Part 3 each subject and each pair of subjects are tested on 25 positions; the test booklets in these two parts had two randomly selected orders to counter possible learning effects.

Each subject's performance on the last 5 positions of the eight practice positions in Part 1 was scored for pairing purposes only. Subjects were paired in score order from highest to lowest.

The pairs were divided into two groups, A and B. Each group had approximately the same number of subjects. The instructions for Part 2 were read and distributed to all subjects. All subjects in group A had higher scores than any subjects in group B on Part 1 of the experiment.

In Part 2 Group A took the test in pairs. Both members of each pair were encouraged to discuss each test stimulus position together thereby discouraging domination by either partner through-out the test set. At the same time Group B was administered Part 3 with no discussion of positions by subjects allowed (i.e. individually).

In the third part of the experiment, Group A took the test individually while Group B took the test in pairs, i.e. the tasks of the two groups were reversed (see Appendix B for Plan of events for pairs experiment).

The first 5 of the 25 positions in Parts 2 and 3 were only for practice purposes while the last twenty positions being considered for scoring purposes.

We chose to let Group A (stronger subjects) take Part 2 of the experiment in pairs first and then Part 3 individually. Since the performance of a subject may degrade (due to fatigue) as the experiment goes on, the strong (>2000) and intermediate (1600-1999) subjects' results were considered more valuable than the results of novice subjects (<1600) whose expected scores were in the 0-4 range. Their performance would effectively serve as a control to the performance of the intermediate and the strong subjects.

Learning factors for subjects may play a role in the experiment. Through the experience of attempting the test stimulus positions subjects may gain knowledge about the test domain which could be reflected by their ability to perform better in the latter part of the experiment. Had we administered Part 2 entirely to individuals (or pairs) and Part 3 entirely to pairs (or individuals) with Part 3 resulting in a higher average score, then we would not have been able to draw any valid conclusions due to the learning factors which may have been involved. Thus, in order to minimize this potential learning factor, we divided the pairs into two groups. Group A participicated in Part 2 of the experiment as pairs and Part 3 as individuals. To counter-balance, this Group B tock Part 2 as individuals and Part 3 as pairs.

Scoring Function

As described earlier performance on the last 5 positions of Part 1 (8 practice positions) was used only for pairing purposes. In Parts 2 and 3, only the last 20 positions were scored. Nearly all the test positions were selected from the point of view that there is only one correct move. In a few exceptional cases there was more than one "Best Move". Subjects were asked to select their "Preferred Move", with optional second, third and fourth choices. If the "Preferred Move" selected by a subject (or pair of subjects) for a given test position was the correct move, one full point was given; if the subject's (pair's) second choice was the correct move, then 1/2 point credit was given; third choice correct was given 1/3 point credit, and fourth choice correct was given 1/4, just as in the earlier Bratko-Kopec experiment. Again subjects were encouraged via the test instructions and verbally to write down as many choices as they may have considered up to four since this could only aid their total scores. However this experiment differed in that now there were 5 practice positions, 10 type T positions and 10 type L positions for scoring purposes; as opposed to the 12 T positions and 12 L positions without practice in the earlier versions of the experiment.

RESULTS OF THE EXPERIMENT

Summary

The main objective of this experiment, to ascertain how the performance of pairs compares with that of individuals, was satisfied, leading to the overall conclusion that a pair of subjects will perform better than either subject can perform alone.

The improvement in pair's scores was mainly due to an improvement in the L-factor of their scores, though throughout the experiment the T-factor also improved slightly. (See Tables 1 and 2 which indicates that L scores for pairs improved over L scores for individuals in each of 5 rating catagories.) Tables 3 and 4 indicate that all pairs, whether low-rated, intermediate, or high-rated tended to benefit from co-operation.

Performance of individuals was consistent throughout with our apriori hypothesis based on previous experimental work. This test format has also proven itself to be a reliable method of measuring chess rating and strength.

We have tested 44 subjects, two of which completed only Parts 1 and 2 (individual tests) of the experiment, leaving us with the complete test results of 42 subjects. The distribution of 44 individual scores on T and L within six rating categories is given in Table 1. The distribution of 21 subject pairs is given in Table 2. The rating allocated for a subject pair is the average of their individual ratings.

Individual Results

Rating Range	Mean T	Mean L	Mean TS	mean 10(T-L)/S	No. of Subjects	S. deviation of TS
1000-1599	1.88	1.29	3.17	0.73	8	2.07
1600-1799	3.25	2.68	5.93	0.48	12	2.39
1800-1999	4.01	4.64	8.65	-0.57	11	2.54
2000-2199	4.40	4.24	8.63	0.14	11	2.45
2200-2399	7.00	8.50	15.50	-15.00	1	0.00
2400 +	8.00	9.00	17.00	-10.00	1	0.00
Overall mean	4.63	4.61	9.24	-1.09	44	

Table 1   Individual's scores

Pair Results

Rating Range	Mean T	Mean L	Mean TS	mean 10(T-L)/S	No. of Subjects	S. deviation of TS
1000-1599	2.67	3.61	6.28	-3.14	3	0.77
1600-1799	3.06	4.04	7.09	-1.09	9	2.35
1800-1999	6.30	4.83	11.14	4.90	3	1.16
2000-2199	6.97	6.90	13.85	0.14	5	0.99
2200-2399	7.00	10.00	17.00	-30.00	1	0.00
Overall mean	5.20	5.88	11.07	-5.84	21	

Table 2   Pair's scores

Note 1 : Two subjects did not do the test in pairs, thus we can only include their individual test results in Table 1.

Note 2 : The number of subjects in Table 2 refers to pairs.

Note 3 : The statistical information list in this report is computed by the Pascal program (on VAX) and STATPAK on MUSIC.

"Mean T" and "Mean L" scores in Tables 1 and 2 are out of ten, and "Mean TS" scores are out of 20. The porportional deviation "(T-L)/S" (computed to determine whether there are differences between performance on type T and type L positions for subjects) is multiplied by 10 for scaling purposes.

Rating range	Improvement on T (%)	Improvement on L (%)	Improvement on TS (%)
1000-1599	42.22%	179.57%	98.24%
1600-1799	-5.98%	50.76%	19.69%
1800-1999	57.01%	4.06%	28.71%
2000-2199	58.54%	62.74%	60.50%
2200-2399	0.00%	17.65%	9.68%

Table 3 Percentage of improvement within each rating category

Rating range	Mean Ind. TS	Mean pair TS	Mean Improved TS
1000-1599	3.17	6.28	3.11
1600-1799	5.93	7.09	1.17
1800-1999	8.65	11.14	2.48
2000-2199	8.63	13.85	5.22
2200-2399	15.50	17.00	1.50

Table 4 Improvement of TS score within each rating category

The following three bar charts illustrate the average performance of individuals and pairs in terms of T scores, L scores and TS scores respectively. The height of the cross-hatched bar graph represents the average of the pairs' scores within each of the six rating categories. The height of the clear bar graph represents the average of the individuals' total scores within each of the six rating categories.

ANALISIS OF VARIANCE FOR THE REGRESSION

Regression analysis (Individual scores) :

We choose rating to be the independent variable, and individual's scores to be the dependent variable.

Intercept = -12.29717

Regression coefficient = 0.01074

Correlation coefficient = 0.744

Standard error of estimate = 2.524

SOURCE OF VARIATION	D.F.	SUM OF SQ.	MEAN SQ.	F VAL
Attributable to regression :	1	332.37	332.37	52.1
Deviation from regression :	42	267.55	6.37	
Total :	43	599.92		

Key : D.F. = degrees of freedom

50

We shall test the null hypothesis that the true value of the regression coefficient is zero, or in other words that there is no relationship between rating and the scores of individuals. The value of the $F(1,42)$ distribution at the 95% confidence interval is 4.08. Since the calculated F values exceeds the critical F value, hat is $F=52.17 >> 4.08$, we fall within the 95% confidence interval that the regression coefficient is greater than 0. Thus, the null hypothesis that there is no relationship tetween ratings and scores on the test can be rejected.

Regression Analysis (Pair scores) :

We choose to let the average rating of a pair of subjects to be the independent variable, and pairs' total scores to be the dependent variable.

Intercept = -16.24391

Regression coefficient = 0.01413

Correlation coefficient = 0.834

Standard error of estimate = 2.201

SOURCE OF VARIATION	D.F.	SUM OF SQ.	MEAN SQ.	F VALUE
Attributable to regression :	1	210.18	210.18	43.39
Deviation from regression :	19	92.01	4.84	
Total :	20	302.19		

Again we shall test the null hypothesis that the true value of the regression coefficient is zero, suggesting that there is no relationship between a pair's average rating and performance. The value of the F(1,19) distribution at the 95% confidence interval is 4.3808. Since the calculated F exceeds the critical F value, that is F=43.39 >> 4.3808, we fall within the 95% confidence interval that the regression coefficient is greater than 0. Thus the null hypothesis can be rejected.

RATING VS SCORES HYPOTHESIS

From the Bratko-Kopec experiment, we composed by extrapolation a rating-score table for this experiment. Given a rating category, one could expect scores to fall within the ranges indicated in the following table.

Rating	Score (TS)
1300 - 1599	0 - 4
1600 - 1799	5 - 6
1800 - 1999	7 - 8
2000 - 2199	9 -12
2200 - 2399	13 -16
2400 +	17 -20

Table 5 Hypothesis ratings vs scores

From the test results listed in Tables 1 and 2, we constructed the following table which shows the effective corresponding number of rating points pairs' scores improve over individuals' scores within each rating catagory.

Rating catagory	Mean improv. in scores (TS)	Mean improvement in Rating
1300 - 1599	2.5 --> 5.5	1300 - 1599 --> 1650
1600 - 1799	6.0 --> 7.0	1600 - 1799 --> 1800
1800 - 1999	8.0 -->11.0	1800 - 1999 --> 2100
2000 - 2199	11.0 -->16.0	2000 - 2199 --> 2350
2200 - 2399	15.0 -->16.5	2200 - 2399 --> 2375 *

* This rating category had only one pair of subjects.

Table 6 Effective pairs improvement in terms of rating.

Example: If the average rating of two subjects falls somewhere in the 1600 to 1799 range, then they are likely to perform like an 1800 rated subject when working as a pair.

Conclusions on the Pairs Experiment

Based on the results of this Pairs Experiment, we can reliably approximate how many rating points a pair of human chess player subjects will gain over their individual rating performance. The improvement in rating performance for pairs within the first rating category ranges from 100 to as much as 350 points, with an average improvement approaching 200 points (Table 6, on previous page).

Discussions with cognitive psychologists led to the suggestion that our experiment may have included built-in bias. That is, pairs' scores may have been superior to individuals' scores simply because one member of the pair was finding the correct move in each or nearly all of the 20 scored test positions, and the other pair member would then comply with the first member's choice. Thus to determine whether pairs' performances were likely to be a result of real cooperation or something else, we re-analyzed the answers of each pair member on the individual test positions (the same position for each person) taking the maximum (most credit given) to form a composite score. If these "hypothetical pairs' scores" proved to be at least as high as the real pairs' scores (see Appendix C) then the above experimental bias could

52

not be disproved.

However, the further analysis of our data satisfactorily discounts such a bias. Given the apriori probability that a pair will score higher than the higher individual score of that pair is 0.5, then the conditional probability of our result, that pairs' scores were higher than individuals' scores in 13 out of 19 cases is computed by

(10/20) / (13/19) = 0.73

which indicates that based on our data, cooperation is likely to occur at least 73% of the time.

In 10 out of 13 cases where pairs' scores were higher than individuals' scores (see Appendix C) the pairs' scores (P) were also higher than the maximum of the individual scores (MP). That is, we retraced through each pair member's performance on the individual test, giving credit for the OK'ed correct answer to any position. The probability of this happening by chance is derived from the computation :

$$p(10) = \binom{13}{10} * (10/20)^{10} * (10/20)^{3}$$

which comes to 0.03, safely within good experimental confidence bounds.

RESULTS OF COMPUTER CHESS PROGRAMS

Following the 1983 World Computer Chess Championship we sent two test sets of 25 positions each, exactly those used the Pairs Experiment described above, to the 22 participants. The test set which (with the exception of 3 positions) was identical to the Bratko-Kopec experiment was labelled "Old Positions" while the other set was labelled "New Positions". The results of 8 computer chess programs on the Old and New Positions are given in Table 7. For each test set the column headed "SCORE" represents the total score of the programs from a maximum of 25. Columns TT and LL represent scores on tactical and lever positions respectively and the sum of these comprises the SCORE. The column "S" represents the total score on the last 20 positions of the test which may be used for purposes of comparison with human results, since for humans the first five positions were considered as practice only. The column "T" represents scores on 10 tactical positions and the column "L" represents scores on 10 lever positions. The sum of columns "T" and "L" comprises column "S". The 8 programs listed (with the exception of AWIT and BELLE) are new to the original Bratko-Kopec test (ibid.). It is noteworthy that all the programs, regardless of their stipulated rating, score relatively highly when compared with humans (Table 4). Most probably these scores do not represent true Bratko-Kopec test scores, but rather the result of its use as a training set. The column labelled 12 * (T-L)/S (the proportional deviation) indicates a unanimous tendency for computer programs to score better on T positions than on L positions (ibid. p64).

Scores on the "New Positions" are rather similar in distribution to those on the original Bratko-Kopec test and generally correspond to rating, particularly within rating categories. On this test set the domination of T over L is not at all as evident as in the earlier test, although the general trend is still apparent in the column 12 * (T-L)/S. This may partially be a result of the positions selected. Based upon the opinions of a number of subjects it is likely that the "New" tactical positions selected were not as clear and distinct from the "New" lever positions as those in the "Old Positions".

Partial credit for 2nd, 3rd, and 4th choices was obtained either directly from the program's output at the end of two minutes of "think" time, or from output of the "Preferred Move" after different periods of time, e.g. 3 minutes, 30 seconds, and 1 minute for 2nd, 3rd, and 4th choices respectively. Although most programs' points scored on the test are derived from preferred moves after two minutes of thought, it would have been more favourable for the experiment that all programs derive their further choices by the same method.

53

COMPUTER RESULTS

OLD POSITIONS

NAME		VCCC RATING	SCORE	TT	LL	S	T	L	10*(T-L)/S
1. CONST'N	(C)	1816	12.00	8.00	4.00	9.00	6.00	3.00	+3.33
2. BEBE	(S)	1885	13.00	10.00	3.00	9.00	7.00	2.00	+5.56
3. PATSOC	(M)	1291	13.00	11.00	2.00	10.00	8.00	2.00	+6.00
4. AWIT	(M)	1660	13.17	6.83	6.33	10.67	5.33	5.33	+0.00
5. BOBBY	(M)	1186	14.00	7.00	7.00	11.00	6.0C	5.00	+0.91
6. PHOENIX	(M)	1780	14.33	7.00	7.33	11.83	6.0C	5.83	+0.14
7. ADV'3.0	(S)	1900	17.00	9.00	8.00	13.50	8.00	5.50	+1.85
8. BELLE	(S)	2200	18.25	11.00	7.25	14.25	9.00	5.25	+2.63
9. MERLIN	(M)	1791	18.50	11.50	7.00	16.00	10.00	6.00	+2.50

NEW POSITIONS

NAME	VCCC RATING	SCORE	TT	LL	S	T	L	10*(T-L)/S
1. SPINKS 4.0	* 1000	4.00	2.00	2.00	4.00	2.00	2.00	+0.00
2. CONST'N	1816	7.08	5.75	1.33	6.83	5.50	1.33	+6.09
3. BOBBY	1186	7.50	4.00	3.50	5.50	3.00	2.50	+0.91
4. PATSOC	1291	8.00	4.00	4.00	7.00	4.00	3.00	+1.43
5. AWIT	1660	9.17	5.83	3.33	7.83	5.50	2.33	+4.05
6. ADV' 3.0	1900	9.83	5.33	4.50	8.00	4.50	3.50	+1.25
7. PHOENIX	1780	11.17	6.00	5.17	9.17	4.00	5.17	-1.28
8. BEBE	1885	12.00	7.00	5.00	10.00	5.00	5.00	+0.00
9. MERLIN	1791	12.33	5.33	7.00	10.00	5.00	5.00	+0.00
10. BELLE	2200	17.92	10.33	7.58	13.67	7.33	6.33	+0.73

TABLE 7

Results of computer programs on the 25 Old Positions (22 from the original Bratko-Kopec test, 3 added) and 25 New Positions. Score is based on a maximum of 25, TT and LL are based on a maximum of 13 and 12 respectively, while S is based on a maximum of 20 (for purposes of comparison with human results) with T and L both out of a maximum of 10 each.

KEY

(S) : program runs on special purpose hardware.
(M) : program runs on a mainframe.
(C) : commercial product.
* : program version runs on TRS-80.

54

Relationship to Computer Performance Over Time

In the Introduction part of this report, we mentioned that there are two methods to increase the computation power, namely: (a) by employing more processors or (b) by switching to a more powerful processor. In this project we have investigated the performance of human chess players working in pairs as compared to a dual processor computer system.

We predicted that computer chess programs will improve their performance on tactical positions if given more time. Since formal lever positions are considered to require specific chess knowledge, more time would not do much to enhance the performance of computer programs on them. We note that the branching factor increases more rapidly as the search gets deeper (bigger search tree). Therefore, we cannot expect the strength of computer chess programs to be doubled as the given time is doubled. On the other hand, human chess players are not able to performa strictly corresponding linearly deeper search when given more time.

TIME SEQUENCE EXPERIMENT

Objective

The purpose of this experiment is to obtain data which might give some insight as to how the performance of humans on sets of problem chess positions varies with time. The humans' performance is also compared on the basis of rating. The performance of computer chess programs on the same test sets over different time allocations is evaluated. The effect of employing two or more computer processors at the same time to attempt to solve these problem positions is also considered and implemented.

Design of the Experiment

A series of six test sets was designed and comprised by three practice positions at two minutes each, four test sets of 10 positions each devised for 30 seconds, one minute, two minutes, and four minutes of solution time, and a test set of 5 positions at eight minutes of solution time each. Thus, the complete experimental sequence consists of 48 positions which when administered with a fifteen minute break period, requires just over two hours. The test sets were administered in a predetermined, not strictly increasing or decreasing sequential time order which varied with each group of subjects tested, e.g. the last group of six subjects were tested in the order practice positions, four minutes, two minutes, fifteen minute break, thirty seconds, one minute, and eight minutes. A total of 18 subjects with Quebec Chess Federation ratings ranging between 1453 and 2358 (1 rated below 1600, 9 rated between 1600 and 2000, 7 rated between 2000 and 2200, and 1 rated over 2300).

Results of the Experiment

The ratings, scores on each test, and total scores (out of a maximum of 45) of each subject are given in Table 8. The data for these subjects ~~when~~ broken down into two groups, namely those rated 1600-1999 (9 subjects) and those rated 2000-2199 (7 subjects) shows one significant trend; that is, on the 4-minute and 8-minute test sets the subjects rated 2000-2199 scored 21.3% and 17.0%, respectively, above the subjects rated 1600-1999, while on the 30-second, 1-minute, and 2-minute test sets the scores of the two groups differed by less than 5% (see Table 9). This result suggests that the stronger chessplayers distinguish themselves over the intermediate ones when given longer periods of time to solve their problem positions. That is, with more time they can develop the deep understanding of a problem position which is necessary to solve it. On the other hand, when given the shorter time test sets (30 seconds, one minute, two minutes) their performance is rather similar to that of the intermediate players. This result, if further substantiated, is rather in line with the findings of Chase and Simon (1973).

TIME SEQUENCE EXPERIMENT

SCORES

NAME	RATING	P (3)	30 SEC (10)	1 MIN (10)	2 MIN (10)	4 MIN (10)	8 MIN (5)	TOTAL (45)
1. Danny Vezeau	1453	0.0	2.50	0.50	4.00	0.50	2.00	9.50
2. Maxime Chauvet	1636	0.0	4.00	5.00	3.83	2.17	2.00	17.00
3. Dave Duchoeny	1793	0.0	4.50	7.00	4.50	4.00	1.25	21.25
4. Michel Barre	1805	0.0	2.00	5.50	3.50	2.92	1.92	15.84
5. Louis Morin	1833	0.0	3.00	6.00	4.50	2.75	3.09	19.34
6. Hugh Brodie	1900	0.5	4.50	5.75	3.33	2.25	1.50	17.33
7. Andy Fletcher	1901	0.0	4.00	5.50	3.00	2.00	3.00	17.50
8. Jean Desforges	1986	0.0	5.83	2.50	4.83	2.33	2.25	17.74
9. Frank Wang	1986	1.0	3.33	4.75	4.50	3.17	4.00	19.75
10. Jean-Francois Martinez	1990	1.0	5.00	5.00	5.00	1.50	3.25	19.75
11. Guy Roy	2050	0.5	4.50	5.00	5.83	3.83	3.25	22.41
12. Steve Boltuc	2080	0.0	4.00	5.00	3.50	7.42	4.00	23.92
13. Luis Jose Alvarez	2041	1.0	2.00	2.75	5.50	3.75	4.00	18.00
14. Daniel Rousseau	2105	1.0	4.25	5.67	3.67	5.50	4.00	23.09
15. Gabor Lorenz	2108	0.5	5.00	5.50	2.00	3.67	1.50	17.67
16. Anthony Ibrahim	2113	0.5	5.00	5.50	1.33	3.75	3.50	19.08
17. Renaud Nadeau	2120	2.0	5.00	9.00	3.50	5.00	3.00	25.50
18. Camille Coudari	2358	1.0	4.50	6.67	7.00	2.50	5.00	25.67

TABLE 8

* Ratings indicated are Quebec Chess Federation (FQE).
* Rating last published in 1976.

Table 9.

TIME SEQUENCE EXPERIMENT

AVERAGE RESULTS FOR INDIVIDUALS

RATING RANGE	30 secs (10 max)		1 min (10 max)		2 min (10 max)		4 min (10 max)		8 min (5 max)		Total (45 max)	
	Ave. Score	Ave. %	Ave. Score	Ave. %	Ave. Score	Ave. %	Ave. Score	Ave. %	Ave. Score	Ave. %	Ave. Score	Ave. %
(Novice) <1600 1 subject	2.50	25.0	0.50	5.0	4.00	40.0	0.50	5.0	2.00	40.0	9.50	21.11
(Intermediate) 1600-1999 9 subjects	4.02	40.2	5.22	52.2	4.11	41.1	2.57	25.7	2.47	49.4	18.39	40.90
(strong) 2000-2199 7 subjects	4.25	42.5	5.49	54.9	3.62	36.2	4.70	47.0	3.32	66.4	21.38	47.5
(Master) >2200 1 subject	4.50	45.0	6.67	66.7	7.00	70.0	2.50	25.0	5.00	100.0	25.67	57.04
Overall Averages	3.82	38.2	4.47	44.7	4.68	46.8	2.57	25.7	3.20	64.0	18.78	41.73
Percent Score Change from Intermediate to Strong		2.3		2.7		- 4.9		21.3		17.0		6.6

AVERAGE improvement in SCORE per 100 rating points: 1.39

REFERENCES

Chase, W.G. & Simon, H.A. (1973) Perception in chess. Cog. Psych., 4, 55-81.

Elo, A. (1978) The Rating Of Chessplayers - Past and Present. Batsford, London.

Draper, N.R. and Smith, H. (1967) Applied Regression Analysis. John Wiley & Sons, Inc., New York.

Groot, A. de (1965) Thought and Choice in Chess. (edited by G.W. Baylor). Mouton, The Hague and Paris. (Translation, with additions, of original Dutch version of 1946.)

Hort, V. and Jansa, V. (1980) The Best Move. RHM Press, New York. (Tranlation, with additions, from original Russian version of 1976.)

Huntsberger, D.V. and Billingsley, P. (1977) Elementarys of Statistical Inference. Allyn and Bacon, Inc., Boston, Mass.

Kopec, D. and Bratko, I. (1982) The Bratko-Kopec Experiment: A Comparison of Human and Computer Performance in Chess. Advances in Computer Chess 3 pp.57-72. (Ed. M.R.B. Clarke), Pergamon Press, Oxford.

Kopec, D., Irazoqui,I. and Bratko, I. (1983) The Updated Brakto-Kopec Test. Computer Chess Digest Annual pp.45-63. (Ed. E. Irazoui), Computer Chess Digest Inc., Huntington.

Lasker, E. (1947) Lasker's Manual of Chess, Dover, New York.

Matanovic, A. (1975) Informator, No. 18. (Center Za unapredivanja saha) Belgrade.

Newborn, M. (1982) OSTRICH/P - a parallel search chess program. Technical Report 82.3, School of Computer Science, McGill University, Montreal, P.Q.

Nievergelt, J. (1977) Information content of chess positions: implications for chess-specific knowledge of chessplayers. SIGART NEWSL. 62, pp. 13-15.

Pachman, L. (1973) Modern Chess Tactics. Routledge & Kegan Paul, London. (Translated by P.H. Clarke from original Czech version of 1970.)

Pritchett, C. and Kopec, D. (1980) Best Games of the Young Grand Masters, Bell and Hyman, London.

Shannon, H.A. (1959) Programming a computer for playing chess. Philos.

58

Mag. 7th Ser., 41, pp. 256-275.

Wade, R. and O'Connell, K. (eds.) (1972) Bobby Fischer's Chess Games. Doubleday and Co., London.

APPENDIX A

MASTER SHEET
19 TACTICAL (T) POSITIONS (Additional)

TYPE		PLAYERS	BEST MOVE	SOURCE	SIDE TO MOVE
T	11.	Kopec - N. Ocipoff	Qxd4	Pan Amer Intcol.74	(W)
T	12.	McKay - Kopec	...Nxf3+	CPHI P.72	(B)
T	13.	Kmoch - Van Scheltinga	Kh3	BCE Diag 55	(W)
T	14.	Kopec - McKay	h7	CPHI P.62 1980	(W)
T	15.	Jansa - Bilek	Na1	TBM #75	(W)
T	16.	Kollberg - Jansa	...Rxf2	TBM #29	(B)
T	17.	Ree - Jansa	...Bc5	TBM #33	(B)
T	18.	Hort - Ribli	Kg2	TBM #96	(W)
T	19.	Hort - Duckstein	Ba6	TBM #78	(W)
T	20.	Kopec - Wagner	Qxf5/Rf6	7 Brd. Bldfld U.Ill '79	(W)
T	21.	Portisch - Huebner	...Ne4+	Best Games P.130, G4	(B)
T	22.	Alekhine - Lasker	Nf5+	MTM P.31 before 25. Nf5+	(W)
T	23.	Bogolyubov - Alekhine	...b4	MTM P.40 before 29...b4	(B)
T	24.	Vasilchuk - Bobolovitch	Nh6	Pachman D27	(W)
T	25.	Weeden - Kleboe	Bxf6	CPHI P.55	(W)
T	26.	Kopec - C. McNab	Bb6+	CPHI P. 167	(W)
T	27.	Hort - Wade	...Kg4	TBM #170	(B)
T	28.	Alekhine - Wolf	Qe3	MTM	(W)
T	29.	Dessor - Hort	...b5	TBM #49	(B)
T	30.	Andersson - Portisch	b4	B.G. p. 100	(W)

T = Tactical position   L = Lever position

Abbreviation for sources:

CPHI = "The Chessplayer's Home Improvement Course"

TBM = "The Best Move"

MTM = "Meet The Masters"

Source	Author(s)
1. The Best Move	Vlastimil Hort & Vlastimil Jansa
2. Modern Chess Tactics	Ludek Pachman
3. Meet the Masters	Max Euve
4. The Chessplayer's Home Improvement Course	Kopec et. al (Pergamon Press, forthcoming)

59