Taken from:
Klaus Marker, marker software, COGPACK, The Cognitive Training Package Manual, Version 7.9, Heidelberg & Ladenburg, August 2007

5. Previous Evaluations

5.1 Clientele

The COGPACK programs were initiated in 1985. They were first tested on outpatients at the Zentralinstitut für Seelische Gesundheit (Central Institute for Mental Health) in Mannheim, Germany which is connected with the University of Heidelberg. Later data came from other centers, such as the Heidelberg University Neurological Clinic and Psychiatric Clinic, as well as the Ulm Rehabilitation Hospital and other indoor- and outdoor facilities. Non-patient groups were provided by hospital personnel and by friends and relations of patients.

Criteria for inclusion or exclusion in the present evaluative material.

Age 16-65 inclusive,
Voluntary participation,
Agreement to storage of personal data
Agreement to storage of ICD number rather than PIN,
No present addiction, no subnormality, no physical or mental disability which would impair use of keyboard/reading of screens,
No seriously acute psychosis, no acute psycho-syndrome at the time of collection,
No acute psychiatric medication (Limit <10 mg Haloperidol or equivalent),
No low-potency neuroleptics or tranquilizers during the day,
Only first runs in the relevant exercise,
Only exercise runs without aids to solution or Adaptive switching,
For normative values: only scores from permutated predetermined series.

In the calculation of Norms (see 5.8) the following were taken into account (as at 30.08.96):

Treatment	(Code)	Persons	Results
Nonpatients	(be 4)	228	1681
Out-patients	(be 3)	169	2466
Day-clinic Patients	(be 2)	120	984
In-patients	(be 1)	390	5836
Total		907	10967

Diagnostic groups	(Code)	Persons	Results
Nonpatients	(grp 1)	228	1681
Schizophrenics	(grp 3)	262	4362
Schizo-affectives	(grp 4)	49	519
Cyclothymics	(grp 5)	57	825
Compulsive Neuroses	(grp 6)	85	1510
Organic Brain Syndromes	(grp 7)	226	2070
Total		907	10967

Evaluations in 5.2 through 5.7 consider:

Treatment	(Code)	Persons	Results
Nonpatients	(be 4)	33	219
Out-Patients	(be 3)	15	125
Day Clinic Patients	(be 2)	76	587
In-Patients	(be 1)	14	71
Total		138	1002

For age and gender, patient and non-patient groups are comparable. For education comparability is limited: in the patients group high school graduation is higher and university graduation is lower than in the nonpatients. The main reason are university drop-outs by illness. Similarly there are persons who completed secondary school but didn't succeed in completing an apprenticeship in the patients group but not in non-patients.

Distribution of Subsamples by Gender

Subsample:	Non-patients		all patients		Total
Gender	f	%	f	%	f
male	19	57.6	59	56.2	78
female	14	42.4	46	43.8	60
Total	33	100.0	105	100.0	138

chi^2= 0.0196, df= 1, p= 0.8886

Age Comparisons for Subsamples

Subsample	M	STD	N
Non-patients	32.12	8.45	33
all patients	30.30	7.65	105

F=1.2200, df1=32, df2=104, p=.2250
t (homVar)=1.1655, df=136, p=.2480. t (heteVar)=1.1063, df=68, p=.2744

Distribution of Subsamples by Education

Subsample:	Non-patients		all patients		Total
Education	f	%	f	%	f
Hauptschule (secondary school)	0	0.0	10	9.8	10
mittlere Reife (school certificate)	5	15.1	12	11.8	17
Lehre (apprenticeship)	10	30.3	40	39.2	50
Abitur (high school graduation)	3	9.1	28	27.4	31
Hochschule (university)	15	45.5	12	11.8	27
Total	33	100.0	102	100.0	135

No information for 3 patients. chi^2=21.807, df=4, p=0.0002

5.2 Acceptability

During the above period only 2 patients totally refused to take part. Partial refusal occurred for two reasons: firstly, storage of some personal data, such as name, diagnosis, or education was refused. Secondly, trainees occasionally asked to end a specific series of exercises because they felt no longer able to concentrate. This was always accepted, though with undesirable consequences: the average scores of the patients improved, as the exercises were only completed by the more capable trainees.

The main cause of initial reluctance was a fear of being unable to cope with the computer keyboard, or of producing poor results. To our own amazement, we hardly ever encountered paranoid attitudes to computers or data storage, although personal data and permission to store it were requested at the very beginning of the process.

When personal data was recorded, there were questions about what medical records meant, who was allowed to inspect them and who were designated as doctors giving further treatment before agreement was given to storage. Occasionally they asked for assurances that only the researchers would see individual data. The offer of a written assurance was refused as „unnecessary„ in all but two cases. About 2/3 of the day-clinic patients agreed to the entry of results in hospital files before carrying out the exercises.

In general, acceptance was high, partly expressed in a desire to repeat chosen exercises, to complete interrupted series at a later time or to take away printed reports, but also in comments reflecting the aims of the training or interest in individual exercises.

Acceptance appeared to be promoted by introductory demonstration of keyboard and mouse, clear exercise and feedback screens, simple input methods, choice of exercise by the trainees themselves. Continuous bad results often led to giving up, occasional bad results were more often an incentive. Complicated input methods (e.g. in ALPHA-BRAVO) are criticized when the number of tasks is high or the trainee is unfamiliar with the keyboard. More monotonous exercise series seem to be preferred more by limited patients, more complicated by the more able. Long series of exercises of similar form were seldom pleasant, but the time spent on the exercises seemed to be of more subjective importance than the number of tasks. In these trial runs several exercises were therefore abbreviated. There was stronger motivation from people directly before beginning training or returning to work, or in real-life test situations like re-acquiring a driving license, especially when patients were unsure of their capabilities. Comparative values were always studied with interest.

5.3 Gvalue-correlations

The file *Correlat.inf* shows the statistical relationships between the Gvalues of the various exercises. It provides the first indications for an evaluative profile: there are important correlations between the exercises included in the table.

between	MEMORY (Alternative 1)	and	EYEWITNESS (Alternative 1),
between	QUANTITY (Alternative 2)	and	LOGIC (Alternative 2),
between	MATHEMATICs (Alternative 1)	and	SEQUENCE (Alternative 1).

Apart from this, correlations so far found have not been high, even when a higher N is occasionally lower than a significant limit. Several less important correlations that have been found reflect on the one hand speed of reaction and on the other similar material, sometimes both (e.g. REACTION and SEARCH).

Notes for the interpretation of tables in *Correlat.inf:*

Pearson Product-Moment Correlations were calculated. The absolute result should be interpreted using the total of value pairs (n) and chance probability (p). Negative correlations result mainly from the way values are calculated: in some Gvalues, low values indicate high performance, e.g. reaction times, in others the same is true of high values, e.g. plus points, cf. 2.1.

5.4 Gvalues and Personal Data

Correlations have been calculated between Gvalues and the following personal data: age, gender, education. Here we are restricting ourselves to one example: EYEWITNESS a. In the table, the Gvalue is the composite value (correct answers in %), GW1 the value for the first, GW2 for the second "split time", more precisely the second half of the task, Age the age of the trainees, Gender coded as 1=male, 2=female, Education coded as in the personal questions in the program from 0=did not complete secondary education to 5=university/college degree, r the coefficient of correlation and p its chance probability.

Correlations for Exercise EYEWITNESS a (N=43)

Variable		GW1	GW2	Age	Gender	Education
Gvalue	r	.89	.86	**-.40**	**-.06**	**.06**
	p	.0001	.0001	**.0245**	**.6891**	**.6819**
GW1	r		.53	-.26	-.22	.02
	p		.0064	.1079	.1679	.9003
GW2	r			-.44	.13	.11
	p			.0153	.5724	.5085
Age	r				.03	.33
	p				.8114	.0481
Gender	r					.07
	p					.6799

The Gvalue does not depend on gender (-.06) or on education (.06); but a significant negative correlation of Gvalue and GW2 with age, which means lower performance when age increases, especially in the 2nd half GW2 where interferences are more important.

The significant correlation with age is an exceptional finding with EYEWITNESS. There are hardly any gender and age correlations with Gvalues of other exercises, thus it seems permissible to evaluate the material without regard to age or gender, at least for a group of 18 through 57-year-olds. For personal data, patients and non-patients were not separated, patients being in the majority. With a more homogeneous sample it is possible to conceive of differences due to age which were here concealed by psychopathological variance.

5.5 Gvalues and other Reference Values

As a check on the production of Gvalues, correlations of Gvalues with simpler performance benchmark values were calculated where the Gvalue of the training program combines many aspects of performance, especially for:

NR	Total of correct answers per run irrespective of time taken.
TS	Average time taken for solution per task irrespective of correctness of answers.

This is mainly relevant for those exercises which produce Gvalues as Total of Plus Points including correctness and time taken, or as average times corrected for errors. Here the present method of producing Gvalues seems acceptable. To show how high the proportion of correctness and speed in a Gvalue is we present an example, exercise COMPARISONS a. In the table TS means Time/Task in sec, Gvalue the combined value (Time/Task, including 20 penalty seconds per error), NR Total correct answers, Age of trainees, Gender (Code 1=male, 2=female), r coefficient of correlation, p its chance probability.

Reference Value Correlations for COMPARISONS a (N= 95)

Variable		TS	NR	Age	Gender
Gvalue	r	.38	-.84	.00	-.03
	p	.0055	.0000	.9474	.7664
TS	r		.19	.11	-.10
	p		.0842	.2939	.6636
NR	r			.06	-.03
	p			.5740	.7864
Age	r				-.14
	p				.1718

The table shows that gender and age do not correlate with any of the values. Combined value Gvalue is decided more by correct solutions NR than by speed TS, which seemed desirable when constructing the program and dictated the relatively high level of time penalty. In the whole group there seemed to be no regular correspondence between strategies for speed and those for accuracy.

5.6 Training Progress in Split Times

Exercise programs calculate split time values in the same way as Gvalues (2.1), but separated for the first and second halves of the total *number* of tasks (GW1, GW2). Some exercises also calculate the number of errors in the first half NF1 (the number of errors for the second half can be calculated as total errors NF -NF1) or record average working times for both halves, details in LEG5*.DBF. Split time comparisons are intended to answer the following questions:

a) Does performance increase in a statistically significant way and to a practically relevant degree even with short training units using COGPACK?

b) Under what conditions (type of task, difficulty. time pressure, stress, personal variables) can one expect an increase in cognitive dysfunctions?

Split times on screen and in Gfiles with notes for interpretation.

Exercises	Screen	Gfile	Structure of split times	Split time useful
ALPHA-BRAVO	no	yes	similar type and difficulty	not known
ANAGRAMS	no	yes	not known	not known
ARCHIVE	no	no	not applicable	hardly
BALL, BORDERS	yes	yes	similar type and difficulty	yes
CALENDAR	no	yes	2nd half more difficult?	not known
CALORIES	yes°	yes	similar type and difficulty	not known
CAR-SYMBOLS	no	yes	similar type and difficulty	not known
CLOCK, COLOR&LABELS	yes	yes	similar type and difficulty	yes
COMPARISONS, COMPASS	yes	yes	similar type and difficulty	yes
CONCEPTS	no	yes	not known	not known
CONFUSION	yes	yes	similar type and difficulty?	probably yes
CONNECT	yes	yes	similar type; difficulty?	not known
DICTATION	no	no	not applicable	no
ELEMENTS	no	yes	similar type and difficulty	not known
EYEWITNESS	yes	yes	similar type and difficulty	yes
FALLINGSTARS	yes	yes	similar type and difficulty	yes
FOLLOW-UP	no	yes	not known	depends on items
GEOGRAPHY	no	yes	similar type. Difficulty?	depends on items
GEOMETRY-polyhedrons	no	yes	similar type. Difficulty?	depends on items
GEOMETRY-triangle,circle,angle	yes	yes	similar type and difficulty	yes
GUESSWORDS	no	yes	not known	not known
INFORMATION	yes	yes	similar type and difficulty	yes
INHABITANTS	yes°	yes	similar type and difficulty	not known
INTERFERENCE a	yes	yes	similar type and difficulty?	probably yes
INTERFERENCE b	no	yes	similar type and difficulty?	probably yes
KEYS	yes	yes	similar type and difficulty	yes
KNOWLEDGE:general,sort, picts, labels	no	yes	not known	not known
KNOWLEDGE:history,literature,movies	yes°°	yes	similar type and difficulty	not known
LABYRINTHS, LINESEGMENTS	yes	yes	similar type and difficulty	yes
LOGIC a,b	no	yes	not known	limited *
LOGIC c..e	no	yes	similar type and difficulty	yes
MATHEMATICS a..j, MATH A, B	yes	yes	similar type and difficulty	yes
MATHEMATICS k..n	yes	yes	similar type; difficulty?	limited *
MONEY	yes	yes	similar type and difficulty	yes
MEMORY	no	no	dissimilar or not applicable	not known
MORSECODE	no	yes	not known	not known
MULTIPLY	yes	yes	similar type and difficulty	yes
NEWorNOT	no	no	probably dissimilar difficulty	not known
NOTE&TONE	yes	yes	similar type and difficulty	yes
NUMBERS a	no	yes	similar type; 2nd half more diff.	not known
NUMBERS b..p	no	yes	similar type and difficulty	not known
ONtheROAD	yes	yes	similar type and difficulty	yes
OPINIONS	yes	yes	similar type; difficulty?	not known
PERCENT, PIECE-WORK	yes	yes	similar type and difficulty	yes
PIESEGMENTS	yes	yes	similar type and difficulty	yes
POEMS	no	yes	not known	hardly
POSITION	yes	yes	similar type and difficulty	yes
PUZZLES	yes	yes	similar type; difficulty?	depends on items
QUANTITY	yes	yes	similar type and difficulty	yes
REACTION	no	yes	depends on settings	partly yes
READING	no	no	not applicable	no
REPRO, ROUTE	yes	yes	similar type and difficulty	yes
SCALES, SCAN	yes	yes	similar type and difficulty	yes
SEARCH	yes	yes	similar type; difficulty?	partly yes
SEQUENCE, STOP	yes	yes	similar type and difficulty	yes
TYPE-IT: new version	no	no	not applicable	no
UFOs, VISUMOTOR	yes	yes	similar type and difficulty	yes
VOCABULARY	no	yes	not known	not known
WISDOM, WHO-OR-WHAT	no	yes	not known	not known

* "limited" means: can only be interpreted taking normative/average values into account.
° only if option "revision allowed" is off
°° only if option "revision allowed" is off and if "Which year?" is asked.

Split *time* values in most cases are split halfs of items or tasks, only *Piece-work* and *Information* use time. Split values comparisons make sense if both parts are similar in type and difficulty and if transfer of learning can be exspected. Here are two examples of Split time comparisons
:

MATHEMATICS Alternative a with familiar low-stress tasks, and EYEWITNESS a which demands speedy reception of several simultaneous stimuli. The following tables show that all patients improve significantly on average in the 2nd split time of MATHEMATICS a. Even schizophrenics achieve this, and have with 18.7 the second highest increase after Neuroses/Personality disorders (27.7), although together with them they have the lowest final. 18.7 points equals one quickly solved task. A more precise analysis shows that even the error quotient goes down from 3.1 to 2.6, and the time taken TS goes down from 16.5 to 14.1 sec/task. The highest theoretical maximum split time score would be, given 0 errors in 20 tasks and 0 sec time/task,= 20*20 + 100/2 starting capital = 450. Maxima achieved were GW1= 431.3 and GW2= 437.1. Group average TS lie between 8.6 and 15.2 sec/task. As with many other exercises we clearly see that non-patients start the first split time with full power and cannot therefore gain much in the second split time. Patients bring down the differential between themselves and non-patients in the second split time.

Split times for all patients MATHEMATICS a

Variable	M	STD
GW1	326.4257	70.8907
GW2	345.8454	59.8802
GW1 - GW2	-19.4197	44.2203

N= 49, t(abh)= 3.0741, df= 48, p= .0160

Split times non-patients MATHEMATICS a

Variable	M	STD
GW1	370.9895	51.6900
GW2	380.2045	51.1650
GW1 - GW2	-9.2150	49.2110

N= 12, t(abh)= 0.6487, df= 11, p= .5357

Split times Schizophrenics MATHEMATICS a

Variable	M	STD
GW1	322.0321	64.4108
GW2	340.7000	64.9288
GW1 - GW2	-18.6679	38.3087

N= 24, t(abh)= 2.3873, df= 23, p= .0441

Split times Cyclothymics MATHEMATICS a

Variable	M	STD
GW1	347.1651	48.7368
GW2	359.2736	56.8907
GW1 - GW2	-12.1085	49.4493

N= 13, t(abh)= 0.8829, df= 12, p= .6014

Split times Neurotics MATHEMATICS a

Variable	M	STD
GW1	317.4612	103.9871
GW2	345.2019	59.8165
GW1 - GW2	-27.7408	51.7074

N= 12, t(abh)= 1.8585, df= 11, p= .1043

Split times Schizophrenics EYEWITNESS a

Variable	M	STD
GW1	61.4737	9.1855
GW2	65.0526	9.8063
GW1 - GW2	-3.5789	

N= 19, t(abh)= 1.7040, df= 18, p= .1178

Tables for EYEWITNESS show: all patients taken as a group improve significantly on average between split times 1 and 2 (lower left table), schizophrenics on their own improve much less and not significantly (upper right table), non-patients equally little and not significantly (lower right table), though starting at a considerably higher level. It is possible that significant results were not achieved because of the small numbers taking part.

Split times all patients EYEWITNESS a

Variable	M	STD
GW1	59.8235	10.8250
GW2	66.5294	10.5061
GW1 - GW2	-6.7059	

N= 34, t(abh)= 3.5205, df= 33, p= .0106

Split times non-patients EYEWITNESS a

Variable	M	STD
GW1	72.2000	8.9169
GW2	75.2000	7.3756
GW1 – GW2	-3.0000	

N= 10. t(abh)= 1.0000, df= 9, p= .6543

Most exercises which have similar tasks (especially of similar difficulty levels) in both halves show, like the examples above, learning gains at least approaching the 5% significance barrier where all patients are taken as a group. Learning gain is often less for non-patients, since they already have good initial values, i.e. begin with equal concentration, and have little room for improvement. If evaluation is diagnosis-specific, schizophrenics often have lower learning gains, in which case diagnosis obviously interacts with the exercise material.

5.7 Prognostic Validity

For occupational rehabilitation purposes the Mannheim Day Clinic, as well as in-house occupational therapy groups and provision in occupational therapy workshops, had the following resources:

> half-day work experience in companies and council offices, limited to a few weeks, pocket-money only, possible during treatment in the Day Clinic.
>
> Rehabilitative work placements: 6-12 months, regular work atmosphere with assistance, wages, possible only after discharge.

Independent of exercises in COGPACK records were kept of occupational therapeutic procedures and used for coding the social workers\lquote assessment below in summary form in code "AVz".

Distribution of the trainees according to degree and nature of success with work experience

Avz	Key	Total	Total
0	No occupational therapeutic target: retired, acutely ill etc.		0
1	Work experience already unsatisfactory in the clinic		27
2	Outside work experience irregularly abandoned		9
	including		
	half-time work experience in companies, council offices	5	
	training, training internships	2	
	regular full-time work placement	2	
3	Outside work placement, regularly ended, with problems		25
	including		
	half-time work experience in companies, council offices	15	
	work experience in their own workplace	3	
	therapeutic workshop	1	
	training, training internships	3	
	part time rehabilitative work placement	1	
	regular full-time work placement	2	
4	Outside work placement, regularly ended, without problems		12
	including		
	half-time work experience in companies, council offices	5	
	work experience in their own workplace	2	
	therapeutic workshop	1	
	training, training internships	1	
	part time rehabilitative work placement	1	
	regular full-time work placement	2	

For research into the relationship between performance in COGPACK and work experience performance, only the following patients were included:

- those with AVz-Code>0 i.e. those who had formulated an occupational target
- fulfilled the criteria set up in 5.1
- who had a record not more than 4 weeks old in one of the following exercise programs: EYEWITNESS a, COMPARISONS a, MATHEMATICS a, REACTION a, STOP d, SEARCH c.

Reasons for the choice of exercises were: COMPARISONS correlated in preliminary investigations with several other performance parameters and seemed to test speedy all-round performance. Many of the work placements were in the offices and sales, which also required basic MATHEMATICS. EYEWITNESS was expected to measure swift orientation in a new and complex situation as well as speedy adaptation to new demands, and thus to predict adaptability to a new work situation. Reaction times with correlations of around .5 seemed useful prognostics for short-term (Zahn & Carpenter 1978) and long-term (Cancro, Sutton, Kerr & Sugerman 1971) outcome, even if there was no specific emphasis on work experience in those papers.

Gvalues for exercises COMPARISONS a, MATHEMATICS a, EYEWITNESS a and REACTION a show significant correlations with success in work placements in the table below. There is no significant correspondence with the Gvalue of the exercise in STOP d and SEARCH c. A binary coded Index AVx, which combines AVz 1 and 2 to 0 and AVz 3 and 4 to 1, produces results which are statistically practically identical using the biserial coefficients.

Correlation of exercise Gvalues with work experience results

		COMPARISONS	MATHEMATICS	EYEWITNESS	REACTION	STOP	SEARCH
Avz	r	-.50	.60	.63	-.63	-.18	.11
	n	51	41	21	18	24	22
	p	.0057	.0032	.0136	.0184	.5971	.6283
AVx	r	-.49	.52	.62	-.56	-.28	-.01
	n	51	41	21	18	24	22
	p	.0059	.0074	.0147	.0326	.1888	.9384

Variance analyses confirm the correlative coefficients. They also show that for SEARCH the AVz-codes 1 and 2 are very underrepresented, thus no decision can be made on the prognostic value of this exercise. For EYEWITNESS, REACTION and STOP a higher N would be desirable, though STOP would hardly gain in prognostic value by this. Only variance analyses for COMPARISONS and MATHEMATICS are given and AVz-codes 1 and 2 are combined because of the low numbers. We see highly significant differences in distribution of the Gvalues in the Avz groups for COMPARISONS and MATHEMATICS.

Gvalues for COMPARISONS in the samples defined above

Work experience	(AVz)	M	STD	N
unsuccessful	(1 or 2)	5.7849	1.6111	30
successful with problems	(3)	4.0932	2.0571	16
successful, no problems	(4)	2.9733	0.9220	5

ANOVA

Source of variance	Sum of squares	df	Estimated variance
between groups	52.1529	2	26.0765
within	142.1544	48	2.9615
total	194.3073	50	

F= 8.8050, p= .0008

Gvalues for MATHEMATICS in the samples defined above

Work experience	(AVz)	M	STD	N
unsuccessful	(1 or 2)	589.1053	128.2813	19
successful with problems	(3)	700.0625	102.6609	16
successful, no problems	(4)	781.3333	11.8940	6

ANOVA

Source of variance	Sum of squares	df	Estimated variance
between groups	209494.1348	2	104747.0674
within	455006.0603	38	11973.8437
total	664500.1951	40	

F= 8.7480, p= .0010

Conclusion: As an aid to the timing of work experience, SEARCH and STOP should be excluded. EYEWITNESS, REACTION, MATHEMATICS, COMPARISONS, can be used, the last two being supported by more data. The distribution of values in the following diagram is first considered separately for each exercise, in order to develop a rationale for the limits in prognosticating success in work experience. In each case there is an overlap for the values of successful and failed work experience. The exact values can often only be seen using higher resolution than is permitted by an A4 diagram.

Distribution of Success in work experience vs. Gvalue in COGPACK exercises

COMPARISONS					MATHEMATICS					EYEWITNESS					REACTION				
Gvalue	f	f	f	f	Gvalue	f	f	f	f	Gvalue	f	f	f	f	Gvalue	f	f	f	f
9.3					823.0			1		82.0	1				1.31	1			
9.1	1		1		807.2			1		80.6					1.28				
8.8					791.3			1	5	79.2					1.25	1			
8.6					775.5	1	1	2		77.9					1.23				
8.4					759.6			2	1	76.5					1.20				
8.2					743.8		1	1		75.1					1.17				
7.9		1			727.9	1		1		73.7			1		1.14				
7.7	1				712.1	1		2		72.4					1.12				
7.5	2	1	1		696.3					71.0					1.09				
7.2	1				680.4		2	1		69.6					1.06				
7.0	2	1	1		664.6			1		68.2			2	1	1.04				
6.8	1				648.7					66.9			1		1.01				
6.5	3	1			632.9	1				65.5	4			1	0.98	1			
6.3	1	1			617.0	1				64.1			1	1	0.95	1			
6.1					601.2					62.7			1		0.93				
5.9					585.3	1		1		61.4					0.90				
5.6					569.5		1			60.0			1		0.87				
5.4	1				553.7			1		58.6	1				0.84				
5.2	1				537.8	3				57.2					0.82	1			
4.9	1	1	1		522.0	1				55.9					0.79				
4.7	1		1		506.1		1			54.5					0.76				
4.5	1			1	490.3					53.1	1				0.73				
4.2					474.4					51.7	1				0.71			1	
4.0	2	1	1		458.6	1				50.4	2				0.68				
3.8		2	2		442.8			1		49.0					0.65				1
3.5					426.9					47.6					0.62		1		
3.3	1			1	411.1					46.2					0.60			4	
3.1			1		395.2					44.9					0.57				
2.8			1	1	379.4	1				43.5					0.54		1		2
2.6			2		363.5					42.1					0.51			1	
2.4			2		347.7					40.7					0.49				
2.1			1	1	331.8					39.4					0.46			1	
1.9	1		1	1	316.0	1				38.0	1				0.43	1			
AVz	1	2	3	4		1	2	3	4		1	2	3	4		1	2	3	4

Gvalue:				
min	1.9183	316.0000	38.0000	0.4325
max	9.3229	823.0000	82.0000	1.3098

Code AVz:
1 Work experience already unsuccessful in clinic.
2 Work experience begun outside, broken off because of poor performance.
3 Work experience outside completed with problems (external assessment).
4 Work experience outside completed without problems (external assessment).

f: Numbers in the cells are frequencies as single digit.

Subprogram COMPARISONS
The area of overlap for the values of successful and unsuccessful work experience is found between Gvalue = 3.42 and Gvalue= 4.91.
If no one is to be excluded who might possibly succeed in work experience outside the clinic, then one should only require a Gvalue < 4.91. The following would then be classified correctly:

Total:	74.5%
AVz 4: External work experience successful, no problems	100.0%
AVz 3: External work experience successful, problems	81.3%
AVz 2: External work experience quit early	66.6%
AVz 1: Work experience in clinic unsatisfactory	71.4%

If uncompleted external work experience is to be avoided as much as possible, then a Gvalue < 3.8, or better < 3.42 should be an entry requirement. With a requirement of Gvalue < 3.42 the following would be classified correctly:

Total:	80.4%
AVz 4: External work experience successful, no problems	80.0%
AVz 3: External work experience successful, problems	50.0%
AVz 2: External work experience quit early	100.0%
AVz 1: Work experience in clinic unsatisfactory	95.2%

If the intention is to classify as few as possible wrongly within groups, the limit will also be set at 3.42. However, this limit would produce too many exclusions of possible successful patients. If the requirement is correct classification for at least all patients with successful work experience and no problems, the limit is < 4.318. The following would then be classified correctly:

Total:	74.5%
AVz 4: External work experience successful, no problems	100.0%
AVz 3: External work experience successful, problems	68.7%
AVz 2: External work experience, quit early	66.6%
AVz 1: Work experience in clinic unsatisfactory	81.0%

Subprogram MATHEMATICS a
The area of overlap for the values of successful and unsuccessful work experience is found between Gvalue = 550 and Gvalue= 743.

If no one is to be excluded who might possibly succeed in work experience outside the clinic, then one should only require a Gvalue > 550. The following would then be classified correctly:

Total:	70.7%
AVz 4: External work experience successful, no problems	100.0%
AVz 3: External work experience successful, problems	93.7%
AVz 2: External work experience quit early	16.7%
AVz 1: Work experience in clinic unsatisfactory	53.8%

If uncompleted external work experience is to be avoided as much as possible, then a Gvalue > 743 should be required. The following would then be classified correctly:

Total:	73.2%
AVz 4: External work experience successful, no problems	100.0%
AVz 3: External work experience successful, problems	43.7%
AVz 2: External work experience quit early	83.3%
AVz 1: Work experience in clinic unsatisfactory	92.3%

If you wish to avoid false classification of as many patients as possible within groups, MATHEMATICS has two equally good limits, 743 as already stated, and a further one at 675, which can be regarded as the mean of both risks of failure. If Gvalue > 675 is required the following would then be classified correctly:

Total:	73.2%
AVz 4: External work experience successful, no problems	100.0%
AVz 3: External work experience successful, problems	75.0%
AVz 2: External work experience quit early	66.7%
AVz 1: Work experience in clinic unsatisfactory	76.9%

Subprogram EYEWITNESS

has an overlap area between 60 and 66. At Gvalue > 60 (% correct answers) all successful work experience (AVz 3 and 4) is correctly classified, but only 60% of those who fail within the clinic (AVz 2 is not represented) and in total reaches a best value of 81.0%. Requiring Gvalue > 66 achieves a correct classification of 76.2%, classifies all failures correctly, but unfortunately misclassifies 50% of those who succeed in external work experience and 43.9% of those who succeed with problems.

Subprogram REACTION

has an overlap area between 0.70 and 0.52. The best overall classification of 77.8% is achieved with a requirement of Gvalue < .60 (error-corrected sec per reaction).

Combined Predictions

It should be possible to improve predictions of success in work experience by using the results of several exercises simultaneously. In the sample we find a satisfactory N only for the combination COMBINATIONS and MATHEMATICS: there is data from both exercises for 28 of the 73 patients. For predictive purposes the apparently optimum limits are used: for MATHEMATICS > 675, for COMPARISONS < 4.318. The following table shows the results.

Distribution of success in work experience vs. combined predictions of MATHEMATICS and COMPARISONS

Avz code	1	2	3	4	Total
Prediction:					
failure in both	5	3	1	0	9
failure in one	4	3	3	0	10
success in both	0	0	6	3	9
Total	9	6	10	3	28

The cells contain simple frequencies. AVz code cf. above.

The requirement for that performance should exceed the defined limits in MATHEMATICS as well as in COMPARISONS as a prerequisite for work experience does classify all failed work experience correctly, all unproblematically successful patients correctly and 6 out of 10= 60% of those who succeeded with problems correctly, a total of 85.7%. This improvement, while not dramatic in percentage terms, convincingly divides the successes from the failures.

It is most difficult to distinguish between Groups AVz 2 (quit external work experience) and AVz 3 (external work experience completed, though with problems). The decision whether to quit or to continue in spite of problems seems not just to be made for performance reasons. there are probably psychosocial factors in play here, where the social abilities of carers and cared-for should interact. Nevertheless the prognostic value of some individual Gvalues is considerable. A survey of extremely divergent values reveals: the 4-week period initially chosen was too long for some patients; for subsiding manias or depressions important changes in performance can occur within four weeks. Therefore there should be only 1-2 weeks between testing the Gvalues and the commencement of work experience.

The correspondence between work experience and Gvalues which we have discovered can be claimed to have a more general validity, as the work experience studied comprises a broad span from clinical occupational therapy through protected workshops to regular full-time work. However, the limits calculated still need confirmation and in particular should not be used for other settings without further investigation.

5.8 Normative Values

Provisional normative values are supplied for some exercise alternatives, mostly Alternative 1. Non-patients and various diagnosis groups are listed separately. Norms are to be found in NORMS.DBF. You can look them up in menu *Norms/Nfile*, and if required print them with menu *File/Booklet of Norms*. They provide a basis for profiles and can be shown after exercise runs as standards of comparison as an alternative to Vfile. The CTOOLS.EXE of the Professional Version can create user defined norm groups and calculate values from the user's own data, which can be selected by menu *File/Paths/Nfile*, cf. App. D. Please mind: 95% of norm data stem from German speaking testees using the German version of COGPACK. So comparability may be assumed for other languages if exercises contain translated items only but must be questioned for those exercises which use different materials like *Reading* or *Poems*. Even if there are identical items e.g. in exercise *Geography/States of the USA* the difficulty might be remarkably different e.g. higher for testees living outside the USA.

The NORMS.DBF supplied contains data calculated at the time of delivery. Criteria for choice and sample characteristics are specified in 5.1. Those fulfilling the sample criteria are included in the calculations. Low numbers have been left in, but you are warned against over-interpretation. In files CKON5-*, parameter minNorms, the user can define minimum Norm-group numbers for use in COGPACK.

It is clear from 5.1 that acute disturbances were excluded, i.e. that values were mostly collected from patients in the rehabilitative phase. Among other reasons, this is because of the general impression that the acuteness of a disease is often more important for performance values than global ICD-diagnoses, though it is still difficult to set acuteness cut-off levels. For organic brain syndromes a broad spectrum of patients from several centers is included, and for a considerable number of these patients we are dealing with minor damage or even recovery at the time the data was collected. Schizophrenic patients included were mainly taken from day-clinic treatment at the beginning of occupational rehabilitation. In contrast, the compulsive neuroses included were more severe and lasting, because of the unavailability of more specialized cut-off points for acuteness and out-patients: consider patients, who in the exercise COMPARISONS show a considerable increase in excitement when having to decide quickly whether two figures are similar or not. Schizoaffective psychoses are listed separately to satisfy the tendency to limit the definition of schizophrenia more strictly. From the figures so far it cannot be decided if this distinction is meaningful for COGPACK. Traces of manic and depressive states in cyclothymia are not distinguished: the reason is insufficient N.

Abbreviations for the provisional Norm-Group data

maximum	or MaxWert	highest numerical value for the group
minimum	or MinWert	lowest numerical value for the group
N		total number of comparative values in the group
Non-patients	or Nichtpatienten	non-patient sample
All patients	or Patienten alle	patients of all diagnoses combined
Schizophrenics	or Schiz. Pt.	schizophrenic patients: subacute or in remission
Schizo-affective	or Schiz-aff. Pt.	schizoaffective patients in the remission phase
Zyklothymia	or Zyklothyme Pt.	endogenic cyclothymic patients, remission phase
Obs.-compulsive	or Zwangsneurosen	atypically severe obsessive-compulsive neurosis
organic brain d.	or Hirnorg. Syndr.	organic brain syndromes of varying derivation

M (Mean), SD (Standard deviation), minimum and maximum refer to combined values (Gvalues). The method of calculating Gvalue is shown on screen: this indicates whether the highest or lowest value is to be regarded as the best score for an exercise.

Exercises were generally presented in series in *Random order*, in order to minimize the effects of sequence and to ensure representativeness. Data deviating from this pattern was included from the following exercises, which mostly depend on education and were chosen freely out of interest by trainees:

SEQUENCE, CALORIES, LOGIC_a, NOTE&TONE, MATHEMATICS, TYPE_IT, NUMBERS, VOCABULARY.

5.9 Performance and Diagnosis

The figures presented cast light mainly on schizophrenics and on exercises for which values taken from permutated series are available in sufficient quantity, since a few extreme values can distort a mean value more the smaller the sub-group N. Since N is still frequently small the following statements should be regarded mainly as demonstrating tendencies. A survey of the Gvalue distribution demonstrates the following results as seen in experimental research:

Schizophrenics almost always show higher scattering of values in comparison with non-patients and cyclothymics. On average, schizophrenics are generally below non-patients, but with the unexpected exceptions: CONCEPTS, REACTION. On average, schizophrenics are sometimes below cyclothymics as well. Whether this occurs seems among other things to depend on the cognitive retardation of the depressives studied. Schizophrenics have the most difficulty keeping up with cyclothymics in LOGIC b, FOLLOW-UP, MATHEMATICS, SEARCH, ANAGRAMS, QUANTITY.

On average, schizophrenics are roughly on a level with cyclothymics for COMPARISONS, REACTION, STOP\-, KEYS, SEQUENCE, PIECEWORK, MEMORY, EYEWITNESS. They do better on CONCEPTS and LABYRINTHS. We also find good scores, sometimes optimum scores, achieved by schizophrenics. It is difficult to perceive a unified group for schizophrenics, and the attempt to define sub-groups is easy to understand.

Split times indicate that schizophrenics take longer to warm up. But compared with non-\-patients the same is true of other diagnostic groups.

The above can be achieved with the normative and split time values supplied. We can add that severe depressions lessen performance and thus Gvalues severely in almost all exercises. This is most noticeable for EYEWITNESS and MEMORY, which otherwise seem useful diagnostic indicators for the after-effects of organic brain disturbances. Performance seems generally to depend more on the degree of acuteness than on global ICD-diagnoses.

5.10 Performance and Medication

The exercises used for test purposes were run without psycho-medication at the time of testing for 9.2% of schizophrenics, 29.4% of cyclothymics, 31.0% of neuroses /per\-sona\-lity disorders. In 5.1 limits are set for neuroleptic medication for acceptance in sample groups, other neuroleptic substance groups were evaluated on the basis of equivalence tables. It is questionable whether the setting of such limits is sensible without regard to individual needs for treatment and tolerance of medication. e.g. the best score in exercise COMPARISONS over 3 years for over 100 patients was omitted because the patient was at the time taking 20 mg Haloperidol per day. In individual cases perceptible improvement under medication was noticeable. No statistical information is yet available, but there are similar reports in e.g. Orzack & Kornetsky 1971, Wohlberg & Kornetsky 1973. We observed other patients whose medication was below the levels set, but whose motor functions were severely limited and slowed: they were therefore excluded from the sample group.

Exercises that require short quick keystrokes or mouseclicks, e.g. SEQUENCE, seems to suffer extrapyramidal neuroleptic side-effects. A typical error: in spite of repeated and comprehended explanation of the repeat function of the keyboard, the finger is not removed quickly enough from the key before the appearance of the next item. As soon as this type of effect was noticed, the exercise was broken off and only resumed after reduction, change or cessation of medication. Exercises were sometimes provided with longer pauses between appearance of items, thus preventing input in the meanwhile, e.g. with COMPARISONS.

The only general rule is that individual tolerance of medication as regards type, dose and acuteness of psychosis must be re-evaluated every time.

5.11 Repetition and Transfer of Learning

Exercise programs should not be expected, as is done in classical test theory for reliability of repetition, to produce the same results when repeated by the same person; one should rather hope that practice will bring swift improvement, a train of argument that is demonstrated above in the case of split time values. Swift improvement is desirable not only for specialist content, but also and even more for general and transferable skills and strategies. In COGPACK transfer of learning of many kinds can be expected, e.g.:

With routines which demand syntax, keyboard, mouse: cues are grasped more securely, repeat keys found more quickly and confidently. Time and correctness improve independent of content.

Comprehension of instructions: e.g. explanations of point values, time deductions and so on are gradually understood to their full extent and only used tactically after repeated feedback.

Patients are often little used to reading or adapting to performance standards set by others. This causes increased warming up times, as demonstrated above in the context of split times.

Learning transferable contents and more general solution strategies.

Such learning transfer occurs not only through repetition of the same exercise and similar items, but also in different exercises done in series. To keep the effects of series-presentation low, i.e. to distribute error variation equally among all exercises, permutation schemes were introduced. Within a single exercise, transfer is assessed using the split time scores.

Repetition of the same exercise has not yet been systematically investigated. For the samples illustrated here, repetition was explicitly excluded. We would expect different answers to this question depending on the exercise and the method of repetition. Preliminary observations of unsystematic repeat runs reveals the following tendencies:

Infrequent repetition of the same exercise at intervals of > 2 months would hardly lead one to expect any transfer, especially for exercises with variable tasks produced by random routines. Here variations in basic state, e.g. acuteness of psychosis, come into consideration.

Frequent repetition at short intervals, e.g. on the same day or directly after one another often produce improvements similar to those observed for split times up to an individually variable best performance, to be followed by a decline because of tiredness, excessive performance pressure etc.

Exercises which produce identical tasks every time cannot be used for repeat testing over short intervals, and are of limited use over longer intervals.

Generally, one should expect tiredness, a decline in concentration and motivation and interference from previous tasks as a result of continuous repetition, particularly when absent-mindedness or depressiveness is a factor.

5.12 Summary

It was possible to show for some of the exercise programs in COGPACK that as regards task-structure, average difficulty and acceptability they are suitable for patients in psychiatric day clinics and rehabilitative establishments: statistically significant and practically relevant training successes occur even in short exercise sequences. Correlations between performance in some exercises and success in subsequent work experience demonstrate significant prognostic value. Correlative statistics point to a dependency on age and education for individual performances, as well as to the possibility of producing profiles from the complete package. Provisional tables of normative values, which are worked into the program itself and its output routines, provide values which are differentiated for non-patients, schizophrenic, schizoaffective and cyclothymic psychoses, severe neuroses and organic brain syndromes, all in the convalescent phase. However, the range of samples needs to be extended. Procedures which are norm-oriented seem hardly adequate for exercises which depend strongly on specific previous experience or are aimed at promoting motivational aspects. On the other hand, in other exercises distinct differences can be seen between average performance and distribution of performance depending directly on diagnosis, though the acuteness of the particular illness appears to be a considerable modifying factor.