Trail Making Test: normative values from 287 normal adult controls

Giovagnoli A.R., Del Pesce M.[1], Mascheroni S., Simoncelli M.[1], Laiacona M.[2], Capitani E.[3]

Istituto Nazionale Neurologico "C. Besta", Milano; [1] Clinica Neurologica, Università di Ancona; [2] Fondazione "S. Maugeri", Centro Medico di Veruno, Divisione di Neurologia, Servizio di Neuropsicologia, Veruno, Novara; [3] Clinica Neurologica III, Università di Milano, Italy

The Trail Making Test (TMT), which explores visual-conceptual and visual-motor tracking, is a frequently used neuropsychological test because of its ease of administration and sensitivity to brain damage. In this paper, norms are provided for the time scores derived from parts A and B, and for the (B-A) difference. The data were collected from 287 adult Italian subjects stratified by gender, schooling and age (from 20 to 79 years). The test scores were affected by age, education and general intelligence (as expressed by Raven's Coloured Progressive Matrices). Only for part A did females have longer time scores than males. Test-retest reliability was high for each score.

Key Words: Trail Making Test — Conceptual and visual-motor tracking — Divided attention — Shifting ability.

Introduction

Since its introduction in 1944, the Trail Making test (TMT) has been widely used as an easily administered test of visual-conceptual and visual-motor tracking, and has become a standard component of many neuropsychological batteries [22]. The test is given in two parts. Part A requires subjects to connect a series of consecutively numbered circles, and thus involves visual scanning, number recognition, numeric sequencing and motor speed. Part B requires subjects to connect a series of numbered and lettered circles, alternating between the two sequences; in this case it is possible to assess mental flexibility in managing more than one stimulus at a time and in shifting the course of an ongoing activity [14]. The difference between the time scores of parts B and A is due to the longer duration and more complex nature of part B in both the adult and child versions [23]. The time difference (B-A) is also considered as reflecting cognitive activity and shifting ability [6]. Neither the number nor the type of errors are sensitive in differentiating brain-damaged patients from healthy subjects [13].
The TMT was originally structured as an intelligence screening test and was only subsequently used to assess brain-injured patients [2]. Reitan's [19, 20] standardization led to the administration format corrently used in adult subjects; a shorter child's version was developed later [21]. The sensitivity of the TMT in detecting the presence of brain damage has been repeatedly confirmed [4, 19, 20, 24, 26, 27]; this sensitivity is mainly due to the fact that the test involves pshychomotor speed and divided attention [20, 26].
TMT performances have generally been found to decline with age [7, 9, 10, 11, 12, 15], the only exception to this being the findings of Boll and Reitan [4]. Intelligence [4, 11, 12, 17, 20] and the degree of education [9, 12, 18] have also been found to affect TMT results. Emotional disturbances, severe anxiety and psychotic symptoms may affect the part B of the test, although rarely to the same extent as brain damage [8].
Despite the widespread use of TMT, no normative data for adult Italian subjects are available, and only a few normative studies have been carried out in other countries [7, 16, 25]. With the aim of filling this gap, we have calculated normative data from 287 normal Italian adults and analyzed the influence of age, education, gender and general intelligence.

Subjects and methods

Subjects

The study involved 287 subjects (154 females and 133 males) without any history of neurological or psychiatric symptoms. Their mean age was 42.2±15.4 years (range 15-79), the mean duration of education was 11.4±4.7 years (range 3-17) and the mean score on Raven's Coloured Progressive Matrices [3] was 24.8±5.1 (range 14-36). Table I shows the distribution of the sample by age and education.

Procedures

We adopted the administration format of Reitan [19, 20]. The examiner first explained the procedure for part A: 25 numbered circles (1.3 cm) printed on a paper sheet (21×29.7 cm) in a standardized random layout had to be connected in the correct sequence using a pencil; the subject was required to complete the test as quickly as

Received 20 January 1996 - Accepted 9 April 1996

TABLE I. *Demographic characteristics of the sample.*

Education	Age <20	20-29	30-39	40-49	50-59	60-69	70-79	
1	0	1	3	14	18	20	7	63
2	8	8	12	9	11	6	2	56
3	0	8	7	11	2	1	0	29
4	3	37	15	12	12	6	0	85
5	0	3	4	1	3	0	1	12
6	0	9	16	5	6	6	0	42
	11	66	57	52	52	39	10	287

Key: 1: primary school, 2: secondary school, 3: trade school, 4: high school, 5: para-university course, 6: university.

possible. A trial run (from number 1 to 8) was first carried out in order to ensure the subject's understanding; only after this trial, was the subject asked to do the complete test (from number 1 to 25). The scoring was based on time alone, taking the number of seconds used to complete the test; the examiner immediately drew attention to any error, which had to be corrected by the subject before he/she was allowed to proceed with the test (thus increasing the time score). The same procedure was followed for part B, which required the subjects to connect alternately the numbers from 1 to 13 and the letters from A to N.

Statistical methods

The TMT scores were analysed by means of simultaneous multiple regression, in order to check the influence of age, education (years of schooling), gender and general intelligence. For each score, a linear regression model was identified that could be used to adjust the original scores for age, education and gender by simply reversing the sign of the parameter. Raven's Coloured Progressive Matrices were not used to adjust the raw scores. Terms were included in the model only when the significance level related to each of them corresponded to p less than or equal to 0.017 (decided on a Bonferroni basis by considering an overall 0.05 level for each measure divided by the three independent variables). The effect of each variable was evaluated within the complete model by partialling out the effect held in common with the other variables.

On the basis of the corresponding regression model, an adjusted score was calculated for each subject by adding or subtracting the contribution of the concomitant variables from the original score. The adjusted scores for each measure were ranked, and the non-parametric tolerance limits for the upper 5% of the values (i.e. the slowest times) were calculated with a 95% confidence. According to Ackermann [1], the outer and inner tolerance limits should be considered separately; the times exceeding the outer limits belong to the worst 5% of the population, and those shorter than the inner limits belong to the best 95% of the population. In both cases, this judgement can be given with a 5% risk of error. For the scores falling between the inner and outer limits, it is not possible to give a controlled judgement, and these can be best defined as borderline values. The use of adjusted test scores may be more informative if they can be standardized

TABLE II. *Scores (mean±SD) for parts A, B and (B-A) of TMT by age, education and gender.*

Age	20	20-29	30-39	40-49	50-59	60-69	70-79
A	32.54 ±9.99	33.45 ±13.03	38.75 ±16.40	48.90 ±23.58	53.83 ±26.30	67.26 ±28.69	84.60 ±23.76
B	63.09 ±16.77	78.06 ±33.69	86.24 ±34.39	111.77 ±53.98	134.50 ±80.06	164.54 ±97.41	336.80 ±197.80
B-A	30.55 ±16.14	44.61 ±27.75	47.49 ±28.15	62.87 ±37.07	80.67 ±63.23	97.28 ±80.03	252.20 ±180.12

Education	1	2	3	4	5	6
A	70.03 ±34.73	44.89 ±15.02	46.07 ±17.46	37.52 ±14.82	42.50 ±21.22	38.69 ±18.90
B	207.84 ±131.02	106.27 ±41.80	97.93 ±38.81	83.61 ±42.60	102.08 ±36.65	74.67 ±25.24
B-A	137.81 ±110.30	61.38 ±33.54	51.86 ±31.70	46.09 ±33.99	59.58 ±25.15	35.98 ±18.73

Key: 1: primary school, 2: secondary school, 3: trade school, 4: high school, 5: para-university course, 6: university.

Gender	Males	Females	Total
A	43.81±23.15	50.38±26.11	47.34±24.96
B	109.33±79.73	122.16±90.55	116.21±95.80
B-A	65.52±65.06	71.77±72.43	68.87±69.06

TABLE III. *The effect of age, education, gender and general intelligence on TMT scores.*

Part A			
Individual regression analysis:	Age	F (1,285)=106.581,	p<0.0001
	Education	F (1,285)=71.024,	p<0.0001
	Gender	F (1,285)=5.017,	p<0.0260
	RCPM	F (1,285)=90.057,	p<0.0001
Simultaneous regression analysis:	Age	F (1,283)=61.638,	p<0.0001
	Education	F (1,283)=22.404,	p<0.0001
	Gender	F (1,283)=4.455,	p=0.0357
	RCPM	F (1,282)=15.143,	p=0.0001
Part B			
Individual regression analysis:	Age	F (1,285)=99.643,	p<0.0001
	Education	F (1,285)=93.244,	p<0.0001
	Gender	F (1,285)=1.598	NS
	RCPM	F (1,285)=89.508	p<0.0001
Simultaneous regression analysis:	Age	F (1,284)=48.556,	p<0.0001
	Education	F (1,284)=43.024,	p<0.0001
	RCPM	F (1,282)=12.717,	p=0.0004
B-A			
Individual regression analysis:	Age	F (1,285)=69.896	p<0.0001
	Education	F (1,285)=74.580	p<0.0001
	Gender	F (1,285)<1	NS
	RCPM	F (1,285)=64.764,	p<0.0001
Simultaneous regression analysis:	Age	F (1,284)=30.890,	p<0.0001
	Education	F (1,284)=35.047,	p<0.0001
	RCPM	F (1,282)=8.095,	p<0.0047

Key: RCPM: Raven's Coloured Progressive Matrices. The effect of RCPM in the simultaneous regression analysis was evaluated within a model that always included age, education and gender.

in some way and so, following the arguments and method by Capitani and Laiacona [5], we have transformed the adjusted scores of parts A and B, as well as the difference (B-A), into Equivalent Scores. A 5-point scale was used, for which 0 indicates a performance corresponding to the worst 5% of the population, and 4 is equal to or better than the median value; the scores 1, 2 and 3 are intermediate between 0 and 4 on a quasi-interval scale.
For each TMT score, test-retest reliability was computed by means of Pearson's correlation coefficient.

Results

The mean scores of parts A and B and (B-A) were, respectively, 47.3±24.91, 116.2±85.8, and 68.9±69.1. Table II shows the mean scores of part A, part B and the (B-A) difference by age, education and gender. The distribution of all of the measures was skewed, with a longer right tail. The correlation between part A and part B was 0.75, that between part A and (B-A) was 0.571, and that between part B and (B-A) was 0.971.
The results of the regression analyses are shown in Table III. The influence of age, education and general intelligence was always significant; only for part A did females show a significantly longer time, but the level of significance was low and did not justify including gender among the adjusting variables. Table IV shows the final regression models of the three TMT measures.
Table V reports the correction grids for the most frequent combinations of age and education; intermediate values can be obtained by means of interpolation or the use of the original linear models shown in Table IV (after reversing the sign of the parameter).
The inner and outer tolerance limits of each measure are given in Table VI.
Table VII shows the Equivalent Score partition of the adjusted scores of the tests and their difference.
A retest was performed by 57 subjects, the test-retest comparisons yielding the following significant correlations: part A: r=0.745, part B: r=0.849, (B-A): r=0.743.

TABLE IV. *Best simultaneous linear regression models of TMT scores.*

Part A
A*=A−0.656 × (age - 42.181) + 1.465 × (education - 11.446)
Part B
B*=B−2.028 × (age - 42.181) + 6.254 × (education - 11.446)
B-A
(B-A)*=(B-A)−1.373 × (age - 42.181) + 4.789 × (education - 11.446)

Key: A*, B* and (B-A)*=adjusted scores; A, B and (B-A)=row scores.
Education is expressed as years of schooling.

TABLE V. *Corrections to be added to or subtracted from the raw scores according to age and education (expressed as years of schooling).*

Part A

	Age											
Education	20	25	30	35	40	45	50	55	60	65	70	75
3	2	−1	−4	−8	−11	−14	−17	−21	−24	−27	−30	−34
5	5	2	−2	−5	−8	−11	−15	−18	−21	−24	−27	−31
8	9	6	3	0	−4	−7	−10	−13	−17	−20	−23	−26
13	17	13	10	7	4	0	−3	−6	−9	−13	−16	−19
17	23	19	16	13	10	6	3	0	−3	−7	−10	−13

Part B

	Age											
Education	20	25	30	35	40	45	50	55	60	65	70	75
3	−8	−18	−28	−38	−48	−59	−69	−79	−89	−99	−109	−119
5	5	−5	−16	−26	−36	−46	−56	−66	−76	−87	−97	−107
8	23	13	3	−7	−17	−27	−37	−48	−58	−68	−78	−88
13	55	45	34	24	14	4	−6	−16	−26	−37	−47	−57
17	80	70	59	49	39	29	19	9	−1	−12	−22	−32

(B-A)

	Age											
Education	20	25	30	35	40	45	50	55	60	65	70	75
3	−10	−17	−24	−31	−37	−44	−51	−58	−65	−72	−79	−85
5	0	−7	−14	−21	−28	−35	−42	−48	−55	−62	−69	−76
8	14	7	0	−7	−14	−20	−27	−34	−41	−48	−55	−61
13	38	31	24	17	10	4	−3	−10	−17	−24	−31	−38
17	57	50	43	36	30	23	16	9	2	−5	−12	−18

Discussion

This study confirmes that performance times significantly increase with age and that education also (and independently) affects both parts A and B, as well as the (B-A) difference. As far as gender is concerned, we only found a marginal advantage for males with part A. On the whole, these results are consistent with the data generally reported in the literature (for a recent review, see Lezak, 14).

TABLE VI. *One-tailed non-parametric tolerance limits for the upper 5% (worse performance) of the adjusted times with 95% confidence. Subjects with scores that are slower than the outer limits are pathological; those with scores that are faster than the inner limits are normal.*

	A	B	(B-A)
Outer limits	≥94	≥283	≥187
Borderline scores	75-93	200-282	144-186
Inner limits	≤74	≤199	≤143

TABLE VII. *Equivalent Scores.*

Equivalent Scores	A	B	(B-A)
0	>93	>282	>186
1	93-69	282-178	186-112
2	68-53	177-136	111-88
3	52-45	135-103	87-58
4	<45	<103	<58

The greater load of part B in comparison with part A can be evaluated otherwise than by simply calculating their difference. For example, we also calculated the time needed to complete part B after adjustement for that needed to complete part A; however, the correlation between (B-A) and (part B adjusted for part A) was very high (0.923), and working with (B-A) is not only definitely more expeditious for clinical use, but also more consistent with the approach generally followed in the literature.

The correlation between part B and (B-A) also deserves comment. The very high correlation (0.971) may be traced back to the fact that the times to complete part B were more than double those required to complete part A, and their greater dispersion was also greater. As a consequence, part B has an a priori much greater weight than part A in determining the difference. From a practical point of view, this suggests that considering (B-A) for diagnostic purposes is not much different from considering part B alone. We think that the additional load of part B can be better accounted for by comparing the results of parts A and B with their respective norms, and subsequently comparing the adjustments deriving from the separate sections of the tasks. We have reported the norms of the (B-A) difference as these are commonly found in the literature, but we suggest looking for whether the Equivalent Scores derived from parts A and B are different. Equivalent Scores allow a diagnostic judgment that partials out the influence of age and education (which we found significantly influenced the task), whereas a simple comparison of the centiles derived from raw scores would be biased as a result of the potential effect of age and education on the discrepancy at issue.

Sommario

Il Trial Making test (TMT), che valuta il modo di procedere in compiti visivi concettuali e visuo-motori, è uno dei test neuropsicologici più frequentemente usati per la sua semplicità di somministrazione e sensibilità nel rilievo di danno cerebrale. In questo articolo sono presentati i dati normativi dei punteggi forniti dalle parti A e B del test e dalla differenza (B-A). La raccolta dei dati è stata effettuata su 287 soggetti adulti Italiani suddivisi in base ad età (compresa tra 20 e 79 anni), scolarità e sesso. L'elaborazione dei dati è stata effettuata in modo da ottenere i limiti di tolleranza ed una standardizzazione con il metodo dei Punteggi Equivalenti. I punteggi del test sono risultati influenzati dall'età, scolarità e capacità intellettiva (espressa dalle Matrici Progressive Colorate di Raven). Le femmine hanno mostrato punteggi lievemente peggiori dei maschi solo nella parte A. Il confronto test-retest ha mostrato un'elevata attendibilità per ogni punteggio del test.

References

[1] ACKERMANN H.: *Mehrdimensionale nicht-parametrische Normbereiche. Metodologische und medi[illegible]ische Aspekte.* Springer Verlag, Berlin, 1985.

[2] ARMITAGE S.: *Analysis of certain psychological [illegible]s used for the evaluation of brain damage.* Pshychol. Mon., 60: 149-156, 1946.

[3] BASSO A., CAPITANI E., LAIACONA M.: *Raven's Coloured Progressive Matrices: normative values on 300 adult normal controls.* Functional Neurology, 2: 189-194, 1987.

[4] BOLL T.M., REITAN R.M.: *Effect of age on performance of the Trail Making Test.* Percept. Mot. Skills., 36: 691-694, 1973.

[5] CAPITANI E., LAIACONA M.: *Aging and psychometric diagnosis of intellectual impairment: some considerations on test scores and their use.* Devep. Neuropsychol., 4: 325-330, 1988.

[6] CORRIGAN J.D., HINKELDEY N.S.: *Relationships between parts A and B of the Trail Making Test.* J. Clin. Psychol., 43: 402-409, 1987.

[7] DAVIES A.D.M.: *The influence of age on the Trail Making Test performance.* J. Cons. Psychol., 24: 96-98, 1968.

[8] GASS C.S., DANIEL S.K.: *Emotional impact on Trail Making Test performance.* Psychol. Rep., 67: 435-438, 1990.

[9] GORDON N.G.: *The Trail Making Test in neuropsychological diagnosis.* J. Clin. Psychol., 28: 167-169, 1972.

[10] GORDON N.G.: *Diagnostic efficiency of the Trail Making Test as a function of cut-off score, diagnosis, and age.* Percept. Mot. Skills., 47: 191-195, 1978.

[11] GOUL W.R., BROWN M.: *Effects of age and intelligence on Trail Making Test performance and validity.* Percept. Mot. Skills 30: 319-326, 1970.

[12] KENNEDY K.J.: *Age effects on Trail Making Test performance.* Percept. Mot. Skills, 52: 671-675, 1981.

[13] KLUSMAN L.E., CRIPE L.I., DODRILL C.B.: *Analysis of errors on the Trail Making Test.* Percept. Mot. Skills, 68: 1199-1204, 1989.

[14] LEZAK M.: *Neuropsychological assessment.* Oxford University Press, New York, 381-384, 1995.

[15] LINDSEY B.A., COPPINGER N.W.: *Age-related deficits in simple capabilities and their consequences for Trail Making performance.* J. Clin. Psychol., 25: 156-159, 1969.

[16] NIELSEN H., KNUDSEN L., DAUGBJERG O.: *Normative data for eight neuropsychological tests based on a Danish sample.* Scand. J. Psychol., 30: 37-45, 1989.

[17] ORGEL S., MCDONALD R.: *An evaluation of the Trail Making Test.* J. Cons. Psychol., 31: 77-78, 1967.

[18] PARSONS O.A., MASLOW H.I., MORRIS F., DENNY P.: *Trail Making Test performance in relation to certain experiments, tests, and subject variables.* Percept. Mot. Skills, 19: 199-206, 1964.

[19] REITAN R.M.: *The relation of the Trail Making Test to organic brain damage.* J. Cons. Psychol., 19: 393-394, 1955.

[20] REITAN R.M.: *Validity of the Trail Making Test as an indication of organic brain damage.* Percept. Mot. Skills, 8: 271-276, 1958.

[21] REITAN R.M.: *Trail making Test for normal and brain-damaged children of early school age.* Percept. Mot. Skills, 33: 575-581, 1971.

[22] REITAN R.M., WOLFSON D.: *The Halstead-Reitan Neuropsychological Test Battery: theory and clinical interpretation.* Neuropsychology Press, Tucson, AZ, 1985.

[23] ROSSINI E.D., KARL M.A.: *The Trail Making Test A and B: a technical note on structural nonequivalence.* Percept. Mot. Skills, 78: 625-626, 1994.

[24] RUSSEL E.W., NEURINGER C., GOLDSTEIN G.: *Assessment of brain damage: A neuropsychological key approach.* John Wiley, New York, 1970.

[25] SELNES O.A., JACOBSON L., MACHADO A.M. et al.: *Normative data for a brief neuropsychological screening battery.* Percept. Mot. Skills, 73: 539-550, 1991.

[26] SPREEN O., BENTON A.L.: *Comparative studies of some psychological tests for cerebral damage.* J. Nerv. Ment. Dis., 140: 323-333, 1965.

[27] WYSOCKI J.J., SWEET J.J.: *Identification of brain-damaged, schizophrenic, and normal medical patients using a brief neuropsychological screening battery.* Inter. J. Clin. Neuropsychol., 7: 40-44, 1985.

Address reprint requests to: Anna Rita Giovagnoli, II Divisione Neurologica, Istituto Nazionale Neurologico "C. Besta", Via Celoria 11, 20133 Milano, Italy.