Experimental Aging Research: An International Journal Devoted to the Scientific Study of the Aging Process
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uear20

Assessing the predictive validity of psychomotor tests as measures of biological age in mice

Donald K. Ingram [a] & Mark A. Reynolds [a]

[a] Molecular Physiology and Genetics Section, Laboratory of Cellular and Molecular Biology, Gerontology Research Center , National Institute on Aging, National Institutes of Health, Francis Scott Key Medical Center , Baltimore, MD, 21224, U.S.A.

Published online: 27 Sep 2007.

To cite this article: Donald K. Ingram & Mark A. Reynolds (1986) Assessing the predictive validity of psychomotor tests as measures of biological age in mice, Experimental Aging Research: An International Journal Devoted to the Scientific Study of the Aging Process, 12:3, 155-162, DOI: 10.1080/03610738608259454

To link to this article: http://dx.doi.org/10.1080/03610738608259454

Experimental Aging Research, Volume 12, Number 3, 1986, ISSN 0361-073x

Assessing the Predictive Validity of Psychomotor Tests as Measures of Biological Age in Mice[1]

DONALD K. INGRAM[2] AND MARK A. REYNOLDS[2]

Two experiments assessed the predictive validity of a psychomotor test battery in male C57BL/6J mice. First, performance was recorded for 66 mice in rotorod, tightrope, grip strength, exploratory activity, and runwheel activity tasks at 24 mo of age. Except in the rotorod task, performance was positively and significantly correlated to lifespan, i.e. better performance indicated longer lifespan. Body weight and body temperature were also significantly correlated with lifespan, while water consumption was negatively related. Using the five behavioral scores in a multiple regression analysis, about 40% of the variance in lifespan was explainable. When measures of body weight, body temperature, and water consumption were added to the regression equation, about 54% of the variance in lifespan could be explained. As revealed by factor analysis, a high degree of interrelationship existed among variables. In a second experiment, 54 mice were tested in the psychomotor battery every 8 weeks from 24 mo. Scores in the tightrope and both locomotor activity tasks revealed age-related declines, whether considering all individuals or only those surviving to 28 mo. Significant correlations between first and subsequent scores indicated stability of individual differences for tightrope and exploratory activity at most ages and for runwheel activity at 26 mo but not later. Rotorod and grip strength scores were not stable and suggested confounding by learning. Significant correlations with lifespan were obtained at some ages for all tests. In contrast to the first experiment, however, there were no significant correlations between lifespan and scores at 24-mo for any test and little correlation among scores. The results demonstrate how the predictive validity of behavioral tests can be assessed but suggest that further refinement of this battery is necessary.

Consistent with the methodology outlined in previous work from our laboratory [14], the present study presents a strategy by which the predictive validity of tests of biological age (BA) can be assessed. Research interest in the concept of BA has concerned the interpretation of variability among individuals of similar chronological age (CA) in the results of various tests designed to assess aging. The central issue is whether inter-individual variability reflects differences in the rate of aging. Although considerable interest has been expressed in developing tests of BA for laboratory animals [13; 24], little agreement exists regarding the methodological strategy to be applied in this effort. Some investigators [7; 19; 26] have emphasized utilizing tests whose scores correlate highly with CA. Others have criticized this approach as being circular [8] or as ignoring an external criterion [23]. If the objective is to apply test scores as more accurate measures of BA than is provided by CA, then why should the correlation of these scores with CA be maximized? Certain tests may yield scores highly correlated with CA but be less related directly to mechanisms underlying aging. At issue in this argument is what should be the external criterion to assess the validity of potential tests of BA [5; 15; 23].

In the previous study from our laboratory, Ingram [14] reviewed concepts and terminology derived from the psychometric literature that might be applied in the development of tests of BA. A suitable criterion for validating a proposed test should be derived from the definition of the construct being assessed. Ingram [14] suggested that a suitable criterion could be derived from the following definition adapted from Comfort [13]:

> Aging is the manifestation of time-related biological processes that result in decreased viability and increased vulnerability of the organism and thus enhance the probability of death.

In agreement with others [4; 20], Ingram [14] suggested that lifespan would be a suitable criterion for validating tests of BA. The ability to predict lifespan would support the validity of a proposed test to reflect differences in viability and vulnerability as demonstrated by differences in the time of death. The *construct* validity of tests of BA might be assessed by their demonstrated ability to differentiate in terms of test scores between two populations with known differences in lifespan. The *predictive* validity of BA tests might be established by demonstrating a correlation between individual test scores and subsequent lifespan. A demonstration of predictive validity would further support the construct validity of the measure.

Significant correlations between lifespan and psychomotor performance have been reported in studies of aged humans [3]. The present study provides a strategy to assess the predictive validity of a psychomotor test battery as applied to aged male C57BL/6J mice. Performance in all tests of the battery was demonstrated previously in our laboratory to be significantly correlated ($rs = 0.40 - 0.72$) with CA in this inbred mouse strain [14]. In addition, individual differences among aged mice were shown to be reliable; that is, temporal stability (retest reliabili-

[1]The Gerontology Research Center is fully accredited by the American Association for Accreditation of Laboratory Animal Care. The authors appreciate the contributions of John Freeman, Edward Spangler, Kathleen Schrieber, and Brian Sievers for data collection; Gunther Baartz, Maurice Zimmerman, and Richard Hiner for design and construction of equipment; Paul Ciesla for computer-aided graphics, and Rita Wolferman for clerical assistance.

[2]From the Molecular Physiology and Genetics Section, Laboratory of Cellular and Molecular Biology, Gerontology Research Center, National Institute on Aging, National Institutes of Health, Francis Scott Key Medical Center, Baltimore, MD 21224, U.S.A.

ty) was demonstrated across a 2-week interval [14]. The present study assessed the correlation between lifespan and test performance at different ages near or beyond the mean lifespan for this mouse strain. This approach is consistent with the life expectancy model for estimating functional age described by Schaie and Parr [23].

The rationale for using aged mice was based on the assumption that variability obtained among aged groups had both statistical and biological significance. Individual differences in psychomotor performance were viewed as generally more reliable among aged mice (26-28 mo) than among young (6-9 mo) mice of this strain in the previous study from our laboratory [14]. In addition, results from another study [17] suggested that statistical indices of psychomotor performance were more frequently significantly correlated with neurotransmitter synthetic enzyme activities in aged (24 mo) mice when compared to young (4 mo) and middle-aged (18 mo) C57BL/6J groups. Thus, we assumed that variability among these aged inbred mice was more reflective of individual differences in the rate of biological aging than among younger mice in which measurement error may be the greatest factor producing variability.

METHODS

Subjects

All subjects ($N = 137$) were male C57BL/6J mice obtained from the colony of the Gerontology Research Center at 23 mo of age. Originally these mice had been purchased at weaning from the Production Department of The Jackson Laboratory (Bar Harbor, ME). The mice were housed in plastic cages with wood shavings as bedding and within the same groups of five throughout their lives. The cages were located in a vivarium maintained at 22 ± 1°C with a 12-hr light: 12-hr dark photocycle with lights on at 6:00 a.m. EST. The mice were provided ad libitum food (NIH-07 formula, 24% protein, 4.2 Kcal/gm) in stainless steel hoppers and water through an automated, filtered system. Fresh cages and bedding were provided weekly. All mice were maintained until death in their cages which were checked daily. The mean lifespan of this strain has been estimated to be 26-27 mo in our laboratory [11].

Procedure

Two experiments were conducted. In Experiment 1, a psychomotor test battery was administered at 24 mo of age. Three birth cohorts (±3 days) were represented. Testing occurred across a 4-mo period. A total of 84 mice were selected for the study, but due to a few early deaths and the loss of data from mechanical failures or scheduling conflicts, the total number presented for data analysis was 66. In Experiment 2, the same test battery was administered at 24 mo of age and then every 8 weeks to surviving mice. Again several birth corhorts were represented and testing was conducted over a 3-mo period. A total of 60 mice began the experiment, but attrition reduced this number to 54 mice for analysis.

Detailed procedures of the psychomotor test battery are described in the earlier report [14]. Tables 1 and 2 provide brief descriptions of each variable and outline the schedule of testing.

Although the schedules differed and procedures varied slightly, every effort was made to standardize general procedures across both experiments. From the beginning of testing in Experiment 1, BW and TEMP were measured weekly throughout life as described previously [21]. TEMP and H_2O data were not collected in Experiment 2. In Experiment 1, all testing was conducted by two experimenters, each concentrating on different tasks. In Experiment 2, another experimenter conducted all tests. Different strain gauges for measuring GS and a different apparatus for testing TR performance were used in each experiment. Procedural differences between experiments involved longer test periods in the EXPL and WHEL tests in Experiment 2 but longer ROTO exposure in Experiment 1.

RESULTS

Experiment 1

Survival. Figure 1 depicts the survival distribution from 25 mo of age for all mice that completed testing. As observed, the median survival was 15.5 weeks, which translated to a median lifespan of about 28.6 mo.

Mean Performance Estimates. Table 1 provides performance estimates for each variable.

Predictive Validity. As presented in Figure 2, the zero-order

TABLE 1

Schedule, Definition, and Mean Performance Estimates of Tests in Experiment 1

Day	Test	Abbreviation	Variable Score	Mean (SEM)
1	Body Weight	BW	grams as measured on electronic balance	35.6 (0.6)
1	Tightrope	TR	mean time in sec before falling from rope over 3 trials	15.9 (1.2)
5	Grip Strength	GS	mean pull in grams on strain gauge over 3 trials	86.1 (1.4)
8 or 9	Rotorod	ROTO	number of falls during 5 min exposure to rod rotating at 3 rpm	25.1 (1.5)
10 or 11	Body Temperature	TEMP	colonic temperature (°C)	36.5 (0.1)
12 or 14	Exploratory Activity	EXPL	number of quarter turns in an oval runway during a 10-min session in the dark	87.1 (4.3)
14 or 16	Runwheel Activity	WHEL	revolutions in a runwheel cage over 48 hr	7,687.9 (594.7)
14 or 16	Water Intake	H_2O	intake of water in grams during exposure to the runwheel cage	20.0 (1.7)

TABLE 2

Schedule, Definition, and Mean Performance Estimates of Tests in Experiment 2

Day	Test	Abbreviation	24 mo		26 mo		28 mo		30 mo	
			n	Mean ± SEM	*n*	Mean ± SEM	*n*	Mean ± SEM	*n*	Mean ± SEM
1	Body Weight	BW*	60	33.7 ± 0.4	31	32.6 ± 0.4	20	31.9 ± 0.6	13	31.5 ± 0.6
1	Exploratory Activity	EXPL*	60	105.5 ± 5.4	42	84.8 ± 2.5	27	61.1 ± 5.9	15	46.6 ± 7.0
2	Rotorod	ROTO*	57	10.0 ± 0.6	42	8.8 ± 2.5	27	4.9 ± 2.5	15	5.5 ± 1.0
3	Tightrope	TR*	58	48.4 ± 3.1	42	30.9 ± 2.6	27	28.4 ± 2.6	15	31.3 ± 7.1
4	Grip strength	GS*	58	55.1 ± 2.0	42	52.1 ± 1.7	76	58.5 ± 2.6	15	60.0 ± 4.3
5-7	Runwheel Activity	WHEL*	44	10,035 ± 1,624	37	9,821 ± 1,625	26	5,645 ± 684	15	4,721 ± 686

*Variable scores are as follows:
BW = grams as measured on electronic balance
EXPL = number of quarter turns in an oval runway during a 15-min session in the dark
ROTO = number of falls during 3-min exposure to rod rotating at 3 rpm
TR = mean time in sec before falling from rope over 3 trials
GS = mean pull in grams on strain gauge over 3 trials
WHEL = revolutions in runwheel cage over 72 hr

correlation coefficients revealed significant relationships between lifespan and all variables recorded at 24 mo, except ROTO scores. Favorable survival was indicated by generally lower water intake and higher body weight, body temperature, and higher scores in the grip strength, tightrope, exploratory activity and wheel activity tests.

A partial correlational analysis was also conducted to control for the influence of BW, which was shown previously to be related to performance in some tests of the psychomotor battery [14]. In the current analysis, BW was also positively related to lifespan. As demonstrated in Figure 2, the magnitude and pattern of the partial correlation coefficients paralleled closely those of their zero-order counterparts; thus, body weight did not confound the correlation between lifespan and test scores.

A separate analysis (results not shown) was conducted to determine the existence of nonlinear (quadratic) correlations between lifespan and the variables. The existence of a significant quadratic component was observed only for the WHEL scores, $R(74) = 0.34$, $p<0.05$, controlling BW.

Factor Analysis. As depicted in Table 3, a high degree of interrelationship among the variables was apparent. The dimensions of these correlations were interpreted by factor analysis of the test scores. Two factor analyses were conducted, the results of which are shown in Table 4. In the first analysis, only scores from the psychomotor tests were submitted; whereas in the second analysis, TEMP and H_2O data were also entered.

When considering only the psychomotor variables, the factor analysis yielded two orthogonally-rotated components. Only the first component was interpreted because it had an Eigenvalue >1.0. An examination of the factor loadings revealed that all variables loaded significantly (>.35) on Factor 1 except ROTO. This component reflects the contribution primarily of locomotor activity – EXPL and WHEL. The results of the second factor analysis yielded two orthogonally rotated components with only the first retained (Eigenvalue>1.0). Again ROTO did not load significantly on Factor 1, whereas all other variables did. Inspection of factor loadings indicated that the two physiological variables, TEMP and H_2O, were closely related to the locomotor component of that factor. Thus, with the exception of ROTO, all variables appeared to be related along a single dimension.

Multiple Regression Analysis. The four psychomotor variables showing significant correlation with lifespan were sub-

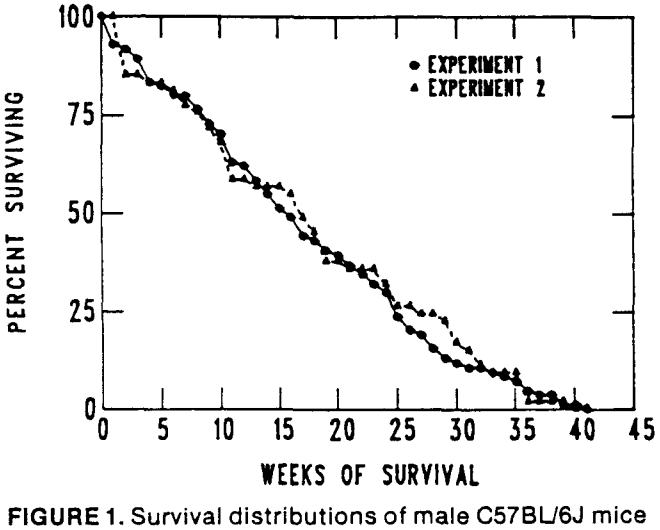

FIGURE 1. Survival distributions of male C57BL/6J mice from Experiments 1 and 2.

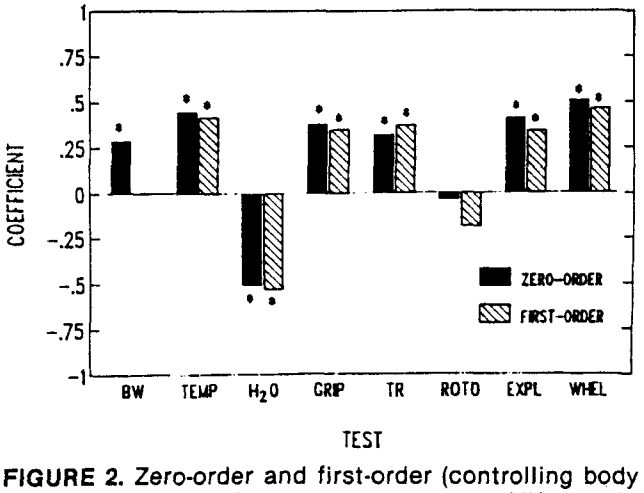

FIGURE 2. Zero-order and first-order (controlling body weight) correlations between test scores and lifespan in 24-mo old male C57BL/6J mice from Experiment 1. *$p<0.05$.

TABLE 3

Intercorrelation Matrix of Test Scores from Experiment 1
(Zero-Order [Pearson Produce-Moment] Correlation Coefficients)

Variable	GS	TR	ROTO	EXPL	WHEL	BW	TEMP
TR	.19						
ROTO	.10	−.25					
EXPL	.22	.20	−.05				
WHEL	.20	.21	−.10	.44***			
BW	.47***	−.17	.51***	.13	.00		
TEMP	.23	.24*	−.04	.47***	.12	.16	
H_2O	−.23*	−.24*	−.05	−.46***	−.36***	−.36***	−.38***

*$p<.05$
***$p<.001$

mitted to a step-wise multiple regression analysis to assess their collective contribution to the prediction of variance in lifespan. Table 5 summarizes the outcome of this analysis. As shown, only two variables, WHEL and GS, entered significantly (F to enter) into the regression equation. These two variables accounted for 34% of the variance (R^2) in lifespan. With all variables entered, the equation was significant ($p<0.001$) and accounted for about 40% of the observed variance in lifespan.

A hierarchical multiple regression analysis was also conducted in which the physiological variables—BW, TEMP, and H_2O—were included in the equation. To control for the contribution of BW in the prediction of lifespan, this variable was entered first into the equation while the remaining variables were entered hierarchically as outlined in Table 6. Three variables—BW, TEMP, and WHEL—entered significantly (F to enter) into the regression equation to account for 49% of the variance (R^2) in the criterion. With all variables entered, the regression equation was significant and accounted for 54% of observed variance in lifespan. Thus, in spite of the interrelatedness of the variables, the inclusion of the physiological variables substantially increased the prediction of the variance in lifespan over that predicted by the psychomotor variables only.

TABLE 4

Summary of Factor Analysis[1] of Test Scores from Experiment 1

	Loading on Behavioral Variables		Loading on All Variables	
	Factor 1	Factor 2	Factor 1	Factor 2
GS	.38	−.14	.37	.11
TR	.33	−.17	.36	−.18
ROTO	−.04	.70	−.01	.73
EXPL	.65	−.04	.74	−.07
WHEL	.64	−.11	.50	−.13
TEMP			.55	−.02
H_2O			−.66	−.05
Variance Explained	67%	33%	75%	25%
Eigenvalue	1.11	0.53	1.82	0.59

[1]Loadings based on Varimax rotation

Experiment 2

Survival. The survival curve for all mice that completed testing in Experiment 2 is presented in Figure 1. As observed, the survival experience of mice in Experiment 2 paralleled that observed in Experiment 1. Median survival was 16.6 weeks, which translated to a median lifespan of 28.9 mo. A Lee-Desu analysis of the survival distributions [18] indicated no evidence of a significant difference between experiments, $D(1)=0.05$, $p>.05$.

Age Effects. Table 2 provides mean performance estimates for each variable at each age (month is equivalent to 4-week intervals). Differences in performance estimates between experiments are noted at 24 mo, but these were likely due to alterations in apparatus and/or procedure.

Figure 3 presents the percent change in mean scores across age for each psychomotor test relative to the scores obtained at 24 mo. As observed in Table 7, the correlations between age and individual scores were significant ($ps<0.05$) for every test except GS. Scores in GS and ROTO exhibited a relative age-related performance increase; whereas, performance in all other tests exhibited an age-related decline by 28 mo that was nearly 50% of the observed mean level at 24 mo. Nevertheless, the relatively low magnitude of the cor-

TABLE 5

Summary of Step-Wise Multiple Regression of Lifespan onto Behavioral Test Scores in Experiment 1

Step	Variable Entered	R	R^2	r	df	F enter	df	F
1	WHEL	.51	.26	.51	1/75	26.41***	1/75	26.41***
2	GS	.59	.34	.38	1/74	9.21***	2/74	19.25***
3	TR	.61	.37	.32	1/73	3.64*	3/73	14.51***
4	EXPL	.63	.40	.41	1/72	2.61	4/72	11.77***

*$p<.10$
***$p<.001$

TABLE 6

Summary of Hierarchical Regression of Lifespan onto all Scores in Experiment 1

Step	Variable Entered	R	R^2	r	df	F enter	df	F
1	BW	.29	.08	.29	1/65	5.83*	1/66	5.83*
2	TEMP	.50	.25	.45	1/64	13.97**	2/64	10.48**
3	WHEL	.70	.49	.56	1/62	16.44**	3/62	14.96**
4	H_2O	.71	.50	−.51	1/61	1.46	5/61	12.35**
5	GS	.72	.51	.38	1/60	1.16	6/60	10.51**
6	TR	.73	.53	.32	1/59	1.99	7/59	9.44**
7	EXPL	.73	.53	.41	1/58	<1.0	8/58	8.17**
8	ROTO	.73	.54	.04	1/57	<1.0	9/57	7.33**

*p≤.05
**p≤.001

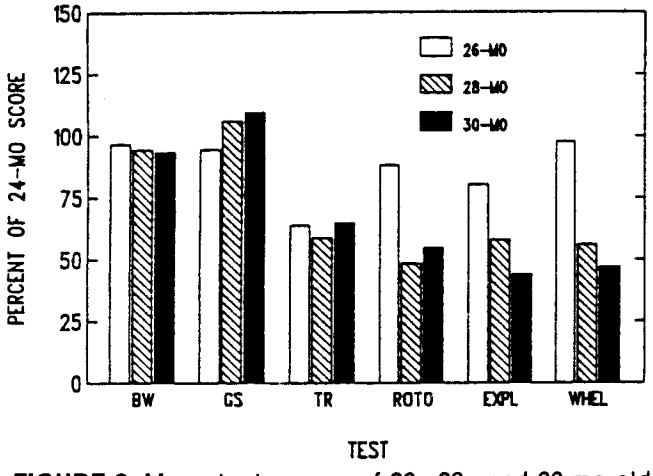

FIGURE 3. Mean test scores of 26-, 28-, and 30-mo old male C57BL/6J mice expressed as percent of score obtained at 24-mo in Experiment 2.

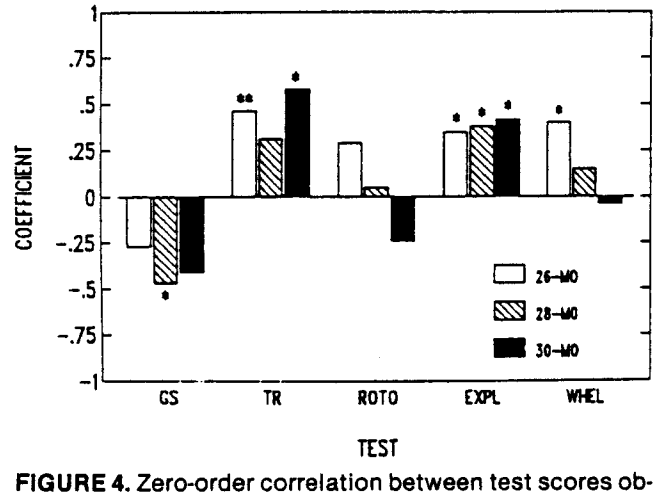

FIGURE 4. Zero-order correlation between test scores obtained at 24-mo and those obtained at 26, 28, and 30 mo in male C57BL/6J mice in Experiment 2. *p<0.05; **p<0.01.

relations between test scores and age (Table 7) indicate a high degree of variability in scores.

Temporal Stability of Individual Scores. Figure 4 presents the correlations between test scores obtained at 24 mo and subsequent scores at older ages. Significant stability (ps<0.05) of the individual scores was apparent for the TR and EXPL tests at several ages and for the WHEL test at 26 mo but not beyond. Animals scoring high in these tests at 24 mo continued to score high at later ages, but the correlation coefficients were only of moderate magnitude (r's = 0.40 - 0.60). In contrast, both the GS and ROTO tests showed poor stability. In fact, high scores at 24 mo appeared to be negatively related to subsequent scores.

Predictive Validity. Figure 5 presents both zero-order and first-order (controlling body weight) correlations between lifespan and test scores at 24, 26, and 28 mo of age. Sample size was considered too small (n<13) for analysis beyond 28 mo of age.

In contrast to the results of Experiment 1, not significant (p<0.05) correlations emerged between lifespan and test scores obtained at 24 mo of age. The correlations for TR and EXPL scores at 24 mo were of similar direction and magnitude as those obtained in Experiment 1, but because of the smaller degrees of freedom in Experiment 2, these coefficients failed to reach statistical significance. Significant correlations (both zero- and first-order coefficents) between lifespan and test performance in the predicted direction were obtained for EXPL at 26 mo and for GS at 28 mo. The zero-order correlation between lifespan and TR scores was significant at 28 mo. In addition, the predicted inverse relationship between ROTO scores (number of falls) and lifespan, which had not been observed at 24 mo in Experiment 1, emerged as a significant correlation

TABLE 7

Correlation of Test Scores with Chronological Age in Experiment 2

Variable	r^1	N^2
GS	.03	125
TR	−.30*	126
ROTO	−.29*	127
EXPL	−.62*	129
WHEL	−.37*	117

*p<.01
[1]Pearson product-moment correlation coefficient
[2]Total number of observations

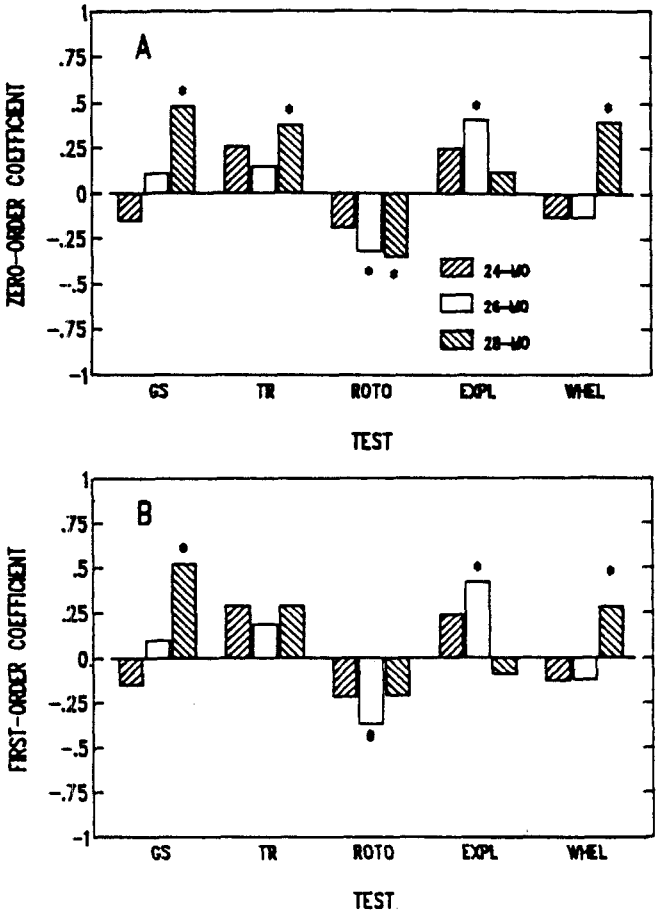

FIGURE 5. Zero-order (A) and first-order (B, controlling body weight) correlations between test scores and lifespan in 24-, 26-, and 28-mo old male C57BL/6J mice from Experiment 2. *$p<0.05$.

(both zero- and first-order coefficients) at 26 mo in Experiment 2. Unlike the strong positive correlation between WHEL scores and lifespan observed in Experiment 1 at 24 mo, no significant correlations emerged until 28 mo in this test when the zero-order coefficient was significant; however, the correction for body weight diminished this relationship.

A separate analysis (results not shown) was conducted to assess for possible quadratic relationships between lifespan and test scores at 24, 26, and 28 mo. Without exception, the quadratic components were not significant ($ps>0.05$) for any variable at any age.

Intercorrelation Among Variables. Smaller sample sizes in Experiment 2 precluded the application of factor analysis of the variables. However, the intercorrelation matrices presented in Table 8 provide an examination of the relationship between any two psychomotor scores obtained at 24, 26, and 28 mo.

In contrast to observations in Experiment 1, few significant ($ps<0.05$) correlations appeared among variables obtained at 24 mo in Experiment 2. WHEL scores were significantly related to EXPL and ROTO scores, but the latter were unrelated. As a general trend, it was evident that the number and magnitude of significant coefficients increased with age beyond 24 mo. By 28 mo, WHEL scores emerged as significantly related to all other variables. ROTO scores were negatively related to all other variables except TR at 28 mo. ROTO scores had been unrelated to other test scores at 24 mo in Experiment 1.

Multiple Regression. As used in Experiment 1, a hierarchial

TABLE 8

Intercorrelation Matrix[1] of Test Scores From Experiment 2

Variable	GS	TR	ROTO	EXPL
24 mo				
TR	.05			
ROTO	−.05	−.25		
EXPL	−.22	.05.	−.08	
WHEL	−.14	.04	−.39**	.29*
26 mo				
TR	−.30			
ROTO	−.34*	−.37*		
EXPL	.05	.22	−.10	
WHEL	−.07	.40**	−.33*	.36*
28 mo				
TR	−.02			
ROTO	−.48**	.10		
EXPL	−.03	.17	−.45*	
WHEL	.39*	.43*	−.49*	.58*

[1]Zero-order (Pearson product-moment) correlation coefficients
*$p<.05$; **$p<.01$

regression analysis applying scores of all tests obtained at 24 mo was conducted to determine their collective contribution to the prediction of variance in lifespan. This analysis yielded a nonsignificant equation, $F(6,37)=1.53$, $p>.05$, accounting for less than 25% of the variance in lifespan. The small sample sizes available beyond 24 mo precluded the application of this analysis for data obtained at other ages.

Predictive Validity of Slopes of Scores. Longitudinal decline in psychomotor test scores could be computed for those individuals surviving from 24 to 28 mo or beyond ($n=26$-27) by estimating the linear slope of scores from each test, except GS, beginning at 24 mo and extending across at least three ages. GS and ROTO were excluded from this analysis because of the apparent age-related increase in performance. Positive correlations between lifespan and slopes were obtained for each variable, which indicated that mice with lower rates of decline in scores exhibited more favorable survival. However, with such small sample sizes, the only significant correlation was for the TR test $r(25)=0.55$, $p<0.01$.

DISCUSSION

Ingram [14] suggested a strategy by which the utility of proposed tests of BA might be assessed. The focus of the strategy is upon statistical evaluation of variability in test scores among individuals of the same CA. The present study was intended to demonstrate in male C57BL/6J mice how the predictive validity of a psychomotor test battery might be evaluated with respect to lifespan as the criterion.

The results of Experiment 1 were encouraging with respect to the predictive value of psychomotor performance among 24-mo old mice. With the exception of ROTO, scores in all other tests showed modest ($rs=0.32$ - 0.51) linear relationships with lifespan. Relatively better performance was associated with better survival experience. WHEL activity was the only variable showing a significant nonlinear (quadratic component) relationship to lifespan. Relatively limited and excessive wheel-running were related to reduced survival. This observation paralleled a previous finding of a quadratic relationship between exploratory activity and lifespan [15]. No such quadratic relationship involving exploratory activity was observed in the present

study, but different methods were used to assess exploratory activity.

The magnitude of the significant linear correlations between test scores and lifespan were virtually unaffected by variation in BW. All the variables, except ROTO, also appeared to be interrelated, as revealed by factor analysis. The two physiological variables, TEMP and H_2O, appeared to be related to this single dimension as they were also related to lifespan.

In spite of the interrelatedness, inclusion of all behavioral and physiological variables into a multiple regression equation could account for over 54% of the variance in lifespan. This R^2 is comparable to that derived from the analysis of the relationship of lifespan to nutritional intake of rats over the lifespan [22], even though the current analysis assessed far fewer variables. Using eight behavioral variables, Botwinick et al. [3] were able to account for about 22% of the variance in survival among a healthy community dwelling population of aged persons.

The analysis in Experiment 2 demonstrated that, with the exception of GS and ROTO, psychomotor performance when assessed repeatedly declines as a function of age in this mouse strain. Even among mice surviving to 28 mo of age, there were longitudinal declines in their performance. These findings supported the previous cross-sectional perspective [14; 16] suggesting significant correlations between psychomotor performance and CA. Moreover, there was evidence of significant stability with age regarding the individual differences in performance. It is interesting that performance at 24 mo would be predictive of relative performance 6 mo later. Performance in the GS and ROTO tests represented exceptions to the observed stability in test scores. Apparently repeated testing, even at 2-mo intervals, permitted confounding by experience.

Regarding the relationship between lifespan and psychomotor performance at 24 mo of age, the results of Experiment 2 failed to replicate those of Experiment 1. No test score obtained at 24 mo was significantly related to lifespan in Experiment 2, although several correlations appeared similar in magnitude and direction to those observed in Experiment 1. Also in contrast to the results of the first experiment was the lack of intercorrelation among the variables at 24 mo. Interrelatedness appeared to increase with age, but the small sample sizes did not permit a summary by factor analysis as accomplished in Experiment 1. Multiple regression, utilizing all behavioral scores at 24 mo to predict lifespan, was also not statistically significant. Again the small sample sizes precluded this type of analysis at older ages.

Although the results of the two experiments regarding performance at 24 mo were not in concert, it was interesting to observe that significant correlations between lifespan and test scores emerged at older ages in Experiment 2. Even ROTO, which had not shown a significant correlation to lifespan at 24 mo in Experiment 1, emerged as a significant correlate at older ages. This finding was notable in view of the relative temporal instability of this test score as mentioned above.

As indicated by the analysis of linear decline (slopes) in performance among mice surviving to at least 28 mo of age, individual differences in the relative change across age appeared to be positively related to survival—the less the rate of decline in performance, the longer the lifespan. Unfortunately, the sample size available for this analysis was relatively low ($n = 26$-27), so that only the TR test yielded a statistically significant correlation.

The lack of agreement between the results of the two experiments regarding performance at 24 mo may stem from undetected methodological differences. When contrasting past results of cross-sectional studies of psychomotor performance in this mouse strain [14; 15; 16] it is clear that differences in methods can produce different results regarding the effects of aging. Thus, although standardization of procedure was emphasized in the current study, sources of inter experiment variation were possible. First, there were differences in procedure and apparatus between experiments. Second, it should be noted that two different groups of technicians conducted the experiments, and no attempt was made to compare their performances (inter experimenter reliability). Third, differences in morbidity and pathology might be another source of inter-experiment variation. The strength of the latter argument, however, is diminished by the observation that the survival curves for the two experiments were similar. The median lifespans were slightly higher than observed previously in this laboratory [11], but this increase might have been due to a selection factor since only mice alive at 25 mo were observed for survival in the present study. Finally, another methodological difference was the experimental histories between the two groups. In Experiment 1, the mice were subjected to the behavioral battery only once at 24 mo and remained untreated except for weekly recordings of body weight and colonic temperature [21]. In Experiment 2, the mice were tested every 8 weeks for as long as they survived. Thus, it is possible that the test experience itself might have influenced the relationships to lifespan, although again no difference in survival experience was apparent between experiments.

Given that individual differences are being assessed in a genetically homogenous population, it is interesting that the variability in scores provided predictive power other than to reflect acute morbidity. The temporal stability of several of the tests, however, rules out acute morbidity as the sole source of individual differences. It is clear from this experiment and others, such as those involving dietary and exercise manipulations [e.g., 10; 12], that environmental factors can profoundly affect the genetic expression of deleterious aging traits even among animals of the same genotype and even at a late age [27]. Phenotypic aging is thus very evident.

In summary, additional analysis is required to provide further confirmation of the predictive validity of this psychomotor test battery. The predictive validity of the tests themselves remains equivocal, but the strategy for their assessment appears logical and productive. Furthermore, it is suggested that this strategy to measure BA can be applied to other noninvasive tests. Methodological improvements are indicated with confirmed standardization of procedure. Further automation of apparatus is planned and might improve the required level of standardization permitting minimum interaction between subject and experimenter. Contamination of the data by differential experience will also have to be further assessed. In addition, the relationship of behavioral performance to specific morbidity and pathology should be evaluated.

It remains the challenge of those interested in developing tests of BA to demonstrate that individual differences representing environmentally- or genetically-associated variation can be proven reliable and valid with respect to predicting a criterion related to the hypothetical construct of aging. Other models of functional age have been suggested which alter the view of aging as a functional decline and emphasize plasticity and diversity [1; 23]. Some investigators have emphasized the multidimensionality of aging that might render fruitless the search for measures of biological or functional age [8; 9]. Because of this multidimensionality, still others have suggested that the construction of a few indices of BA is not an obtainable goal but that the use of multiple tests to develop individual profiles across a range of functional tests may prove to be the most useful approach [2, 5]. What remains, though, is the practical need and

the research demand to search for useful instruments [7; 9; 24; 25]. The challenge of this effort will be to validate proposed tests. The use of test batteries as purely descriptive tools of aging will prove much less productive than the demonstrated application of such batteries for a particular objective. The prediction of lifespan is clearly a desirable application. This demonstrated utility can then be applied to the assessment of treatments which purport to alter aging rate.

REFERENCES

1. Baltes, P.B. Functional age and social policy in aging. In J. Birren et al. (Eds.), *Aging: A challenge to science and society,* Vol. 3, *Behavioral Sciences.* Oxford: Oxford University Press, 1983, pp. 435-438.

2. Borken, G.A., & Norris, A.H. Assessment of biological age using a profile of physical parameters. *Journal of Gerontology,* 1980, *35,* 177-184.

3. Botwinick, J., West, R., & Storandt, M., Predicting death from behavioral test performance, *Journal of Gerontology,* 1983, *33,* 755-762.

4. Brown, K.S., & Forbes, W.F. Concerning the estimation of biological age. *Gerontology,* 1976, *22,* 428-437.

5. Chown, S. Profiles of abilities. In J. Birren et al. (Eds.), *Aging: A challenge to science and society,* Vol. 3, *Behavioral sciences.* Oxford: Oxford University Press, 1983, pp. 264-275.

6. Comfort, A. *Ageing: The biology of senescence.* London: Routledge and Kegan Paul, 1964.

7. Comfort, A. *The biology of senescence.* New York: Elsevier, 1979.

8. Costa, P.T., Jr., & McCrae, R.R. Functional age: A conceptual and empirical critique. In S.G. Haynes & M. Feinleib (Eds.), *Epidemiology of aging.* NIH Pub. No. 80-969. Washington: U.S. Government Printing Office, 1980, pp. 23-46.

9. Davies, A.D.M. Is there a need to estimate functional age. In J. Birren et al. (Eds.), *Aging: A challenge to science and society,* Vol. 3, *Behavioral sciences.* Oxford: Oxford University Press, 1983, pp. 346-360.

10. Goodrick, C.L. The effects of exercise on longevity and behavior of hybrid mice which differ in coat color. *Journal of Gerontology,* 1974, *29,* 129-133.

11. Goodrick, C.L. Life-span and the inheritance of longevity of inbred mice. *Journal of Gerontology,* 1975, *30,* 257-263.

12. Goodrick, C.L. Body weight increment and length of life: The effect of genetic constitution and dietary protein. *Journal of Gerontology,* 1978, *33,* 184-190.

13. Harrison, D.E. Experience with developing assays of physiological age. In M.E. Reff & E.L. Schneider (Eds.), *Biological markers of aging.* NIH Pub. No. 82-2221. Washington: U.S. Government Printing Office, 1982, pp. 2-12.

14. Ingram, D.K. Toward the behavioral assessment of biological aging in the laboratory mouse: Concepts, terminology, and objectives. *Experimental Aging Research,* 1983, *9,* 225-238.

15. Ingram, D.K., Archer, J.R., Harrison, D.E., & Reynolds, M.A. Physiological and behavioral correlates of lifespan in aged C57BL/6J mice. *Experimental Gerontology,* 1982, *17,* 295-303.

16. Ingram, D.K., London, E.D., Reynolds, M.A., Waller, S.B., & Goodrick, C.L. Differential effects of age on motor performance in two mouse strains. *Neurobiology of Aging,* 1981, *2,* 221-227.

17. Ingram, D.K., London, E.D., Waller, S.B., & Reynolds, M.A. Age-dependent correlation of motor performance with neurotransmitter synthetic enzyme activities in mice. *Behavioral and Neural Biology,* 1983, *39,* 284-298.

18. Lee, E., & Desu, M.A. A computer program for comparing K samples with right-censored data. *Computer Programs in Biomedicine,* 1972, *2,* 315-321.

19. Ludwig, F.C., & Smoke, M.E. The measurement of biological age. *Experimental Aging Research,* 1980, *6,* 497-522.

20. Nuttal, R.L. The strategy of functional age research. *Aging and Human Development,* 1972, *3,* 145-148.

21. Reynolds, M.A., Ingram, D.K., & Talan, M. Relationship of body temperature stability to mortality in aging mice. *Mechanisms of Ageing and Development,* 1985, *30,* 143-152.

22. Ross, M.H., Lustbader, E., & Bras, G. Dietary practices and growth responses as predictors of longevity. *Nature,* 1976, *262,* 548-553.

23. Schaie, K.W., & Parr, J. Concepts and criteria for functional age. In J. Birren et al. (Eds.), *Aging: A challenge to science and society,* Vol. 3, *Behavioural sciences.* Oxford: Oxford University Press, 1983, pp. 249-263.

24. Schneider, E.L., Reff, M.E., Finch, C.E., & Weksler, M. Potential application of biological markers for assessing interventions of physiological aging. In M.E. Reff & E.L. Schneider (Eds.), *Biological markers of aging.* NIH Pub. No. 82-2221. Washington: U.S. Government Printing Office, 1982, pp. 237-240.

25. Shock, N.W. Indices of functional age. In D. Dannon, N.W. Shock, & M. Marois (Eds.), *Aging: A challenge to science and society,* Vol. 1, Oxford: Oxford University Press, 1981, pp. 270-286.

26. Skalicky, M., Hofecker, G., Kment, A., & Niedermuller, H. Models of the biological age of the rat: II. Multiple regression models in the study of influencing aging. *Mechanisms of Ageing and Development,* 1980, *14,* 361-377.

27. Weindruch, R., & Walford, R.L. Dietary restriction in mice beginning at 1 year of age: Effect on life-span and spontaneous tumor incidence. *Science,* 1982, *215,* 1415-1418.