

6639
TILT 11

(3)

A note on "Familiar size and the perception of depth" 235
 BY WILLIAM H. ITTELSON

Controlling interaction in small group research 241
 BY MARTIN GROSSACK

Cues, expectations, and first impressions 245
 BY MARTIN GROSSACK

Time perspective and feeling-tone: A study in the perception of the days 253
 BY MAURICE L. FARRER

Simple auditory reaction time in aged adults 259
 BY WALTER D. OBRIST

Atypical tonal memory 267
 BY CYRIL C. O'BRIEN

Methodological considerations underlying electrodermal measurement 271
 BY WILLIAM W. GRINGS

Factors influencing routine performance under noise: I. The influence of "set" 283
 BY E. VICTOR MECH

Certain factors in producing complexity of response to a single pulse of light 299
 BY S. HOWARD BARTLEY AND FRANK RAY WILKINSON

Quantitative studies of verbal reinforcement in classroom situations. I. Differential reinforcement related to the frequency of error and correct responses 307
 BY DONAVON AUBLE AND E. VICTOR MECH

Hypnosis as emotional response: A theoretical approach 313
 BY HENRY GUZE

Problem-solving behavior in men with frontal or occipital brain injuries 329
 BY W. S. BATTERSBY, H. L. TEUBER, AND M. B. BENDER

The performance of albino mice in the maze situation with the utilization of light as motivation and its relative absence as reinforcement 333
 BY WILLARD E. CALDWELL AND H. LYNN WOMACK

Group memory of meaningful material 361
 BY HOWARD V. PERLMUTTER

Philosophic assumptions in operational psychology 371
 BY EDWARD NEWBURY

Published as a separate and in *The Journal of Psychology*, 1953, **36**, 169-179.

THE INTERCORRELATIONS BETWEEN MEASURES OF SCHOOL LEARNING*

Department of Education, Yale University

J. W. TILTON

A. THE PROBLEM

This article supplements an earlier one (7). In the earlier article protest was made against the use of poor measures of learning as criteria for testing the concept of intelligence as a general ability to learn. The measures of learning used have been very unreliable. Furthermore, no thought seems to have been given to the possibility that there was relatively little opportunity for the ablest to profit from the learning experience, nor to the possibility that the tests used could not register what the ablest did learn. With these considerations in mind, measures of learning were described in the earlier article which correlated highly enough with intelligence test scores to be consistent with the idea of intelligence as ability to learn.

At that time, no attempt was made to examine the generalness of learning. The present article describes such an examination of learnings in a school situation. The report consists of five parts: (a) a description of the data to be used as to their adequacy along two of the three lines mentioned above; (b) a description of some further requirements when the generalness of learning is in question; (c) presentation of data as to the generalness of learning in one school situation at the fourth and fifth grade level; (d) a comparison of the factor pattern at the fourth and fifth grade level with the factor pattern of prior gains; and (e) some observations concerning the predictability of improvements in score at the fourth and fifth grade level.

B. THE DATA

The tests used were the Unit Scales of Attainment. The initial testing was done with the Division I battery which is designed for Grades 3 and 4. Forms A and B were administered in the fall to the fourth grades used by the New Haven State Teachers College for teacher-preparation purposes. The end tests were Forms A and B of the Division II battery designed for Grades 5 and 6. They were administered in the spring about 20 months after the initial testing.

*Received in the Editorial Office on October 5, 1952, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

TABLE 1
RELIABILITIES

	Initial scores	Final scores	Improvements
Reading	.791	.856	.364
Social Studies	.838	.950	.694
Literature	.831	.834	.521
Science	.825	.768	.360
Arithmetic Problems	.855	.910	.629
Arithmetic Fundamentals	.852	.907	.769
Spelling	.964	.934	.692
English	.889	.893	.744
Averages	.85	.88	.60

Two features of this procedure were designed to give better reliabilities than are ordinarily obtained, two-year gains and two-form measures. The reliabilities shown in Table 1 are Form A with Form B coefficients stepped up with the Spearman-Brown formula for the Form A and Form B combined measures.

Were the tests valid enough to provide full measures of the learning of all pupils?

It might be claimed that no standardized achievement test battery measures the outcomes of enriched teaching for the ablest students. That may be the case. But so far as standardized achievement testing can take care of the situation, there was plenty of opportunity in the end tests for the ablest students to make high scores. The Division II test range of difficulty did not impose any obvious restrictions upon the measures of improvement.

C. SPECIALIZING INFLUENCES

The measures are fairly reliable. There is plenty of "top" in all the tests. If general improvement is not found, can a general learning ability be denied? By no means. Specializing influences could have counteracted the influence of the general learning ability.

Suppose the group were composed of upper-class college students with a variety of major interests and goals, pre-medical, pre-engineering, etc. The classics majors would make the largest gains in language, the lowest gains in science and mathematics. Would low or even negative correlations between gains mean that there is no general ability to learn? How far up the educational ladder do pupils go without specialization of interest and application? Possibly specialization is present from the start, gradually reducing the evidence for a general ability. If that is the case, the fourth and fifth grades are a good level at which to secure data; a compromise between the lower

level at which so few "subjects" are taught and the upper level at which greater specialization is to be expected.

There are, however, other specializing influences. There are specializing influences in the teaching pattern. It was pointed out in the earlier article that teaching may provide plenty of opportunity for all to learn or it may cause the ablest pupils to mark time while those of lesser ability are helped to make large gains. The only safeguard against this, in the present case, is the probability that the teachers were better than average teachers, and that more than usually is the case, each child was given opportunity and help according to his ability.

But can it be assumed that the 10 teachers involved were alike in their ideas of the relative importance of the eight areas measured, alike in their relative interest in these areas, and alike in their relative abilities to teach the areas effectively? Perhaps the largest gains in reading were to be had only by being in the class taught by Miss A, and perhaps the smallest gains in arithmetic were made by the pupils taught by Miss Y. Under such circumstances, lack of intercorrelation between reading and arithmetic in the combined group would but reflect the heterogeneity of the teaching patterns and would be of little worth as evidence against the existence of a general ability to learn.

In order to eliminate such heterogeneity, one should collect data for many classes so that classes could be dropped freely without serious loss of reliability in the resulting homogeneous group. In the present case, data were available for only eight different instructional groups. Very little elimination was possible. Only two groups were eliminated, and this was done crudely by inspection. They were the two groups which exhibited the most unique pattern of gains. Group A made obviously large gains in social studies, literature, science, and arithmetic, but made low gains in English. Group B, on the other hand, made high gains in English but low gains in arithmetic fundamentals. The intercorrelations for the total group ($N=162$) and for the reduced group ($N=102$) after Groups A and B were eliminated are shown in the first two columns of Table 2. As was to be expected, the correlations between social studies, literature and science were decreased by the elimination of the two groups. The effect upon the correlation between arithmetic problems and arithmetic fundamentals could not be predicted because the elimination of Group A would tend to lower it while the elimination of Group B would tend to raise it. The correlations of gains in these five areas with gains in English were raised as was expected. The over-all average change was slight but the intercorrelations for $N=102$ are sub-

TABLE 2
INTERCORRELATIONS BETWEEN GAINS

	N=162	N=102	Boys N=48	Girls N=54	Boys and girls r's averaged	N=102 corrected for attenuation
R with <i>S</i> ₀	.216	.187	.225	.209	.217	.372
<i>I</i>	.143	.154	.036	.228	.132	.354
<i>Sc</i>	.234	.221	.240	.202	.221	.609
<i>AP</i>	.230	.316	.332	.314	.323	.660
<i>AF</i>	.210	.266	.304	.237	.270	.502
<i>Sp</i>	.073	.065	.145	.021	.129	.083
<i>E</i>	.227	.237	.317	.173	.245	.456
<i>S</i> ₀ with <i>I</i>	.699	.590	.580	.744	.662	.981
<i>Sc</i>	.676	.314	.368	.334	.351	.629
<i>AP</i>	.407	.430	.441	.459	.450	.651
<i>AF</i>	.369	.408	.399	.480	.439	.564
<i>Sp</i>	.258	.164	.253	.125	.189	.237
<i>E</i>	.311	.400	.406	.411	.409	.556
<i>I</i> with <i>Sc</i>	.607	.233	.254	.239	.246	.538
<i>AP</i>	.361	.319	.265	.393	.329	.557
<i>AF</i>	.406	.426	.279	.525	.402	.672
<i>Sp</i>	.194	.171	.122	.194	.158	.285
<i>E</i>	.192	.334	.293	.386	.339	.537
<i>Sc</i> with <i>AP</i>	.383	.325	.537	.115	.326	.683
<i>AF</i>	.286	.142	.244	.066	.155	.270
<i>Sp</i>	.202	.171	.150	.186	.168	.342
<i>E</i>	.121	.268	.320	.234	.277	.518
<i>AP</i> with <i>AF</i>	.478	.536	.504	.590	.547	.770
<i>Sp</i>	.213	.251	.303	.236	.270	.381
<i>E</i>	.255	.366	.375	.354	.365	.535
<i>AF</i> with <i>Sp</i>	.303	.227	.123	.292	.208	.311
<i>E</i>	.380	.527	.481	.579	.530	.697
<i>Sp</i> with <i>E</i>	.389	.345	.205	.462	.334	.481
Averages	.315	.300	.304	.314	.309	.510

mitted as a more valid reflection of ability to learn because of some elimination of what is interpreted as heterogeneity of teaching pattern.

In Columns 3 and 4, the intercorrelations are shown for boys and for girls. They are shown separately for two reasons. One is to show that in spite of the reduction from 162 to 102, there is enough reliability left in the data and enough positive intercorrelation to have prevented the occurrence of a single negative correlation. The other reason for the separation was to test the influence of the specializing influence of sex upon the intercorrelations reported in Column 2. It was predicted that the average *r*'s in Column 5 would tend to be larger than the *r*'s in the second column. They did tend to be larger, and the average difference was 3.2 times its standard error.

Other things being equal, intercorrelations for boys or for girls would be better data for the present purpose; but because of the greater reliability of the *r*'s in Column 2 for boys and girls, the Column 2 *r*'s are submitted as the best data available under the circumstances. They are shown in the last column of Table 2 corrected for attenuation.

D. EVIDENCE OF A GENERAL ABILITY TO LEARN

The absence of negative coefficients in Table 2 is itself evidence of the generality of learning in the eight areas. The substantial loadings with Factor 1 shown in Table 3 are as good evidence of a general ability to learn

TABLE 3
FACTOR LOADINGS BEFORE ROTATION FOR THE EIGHT MEASURES OF IMPROVEMENT, BASED ON THE CORRECTED *r*'S IN THE LAST COLUMN OF TABLE 2

Area	Factor 1	Factor 2	Factor 3	Factor 4	<i>r</i> ²
R	+.63	-.18	+.46	-.10	.65
<i>S</i> ₀	+.84	+.52	-.09	-.07	.99
<i>I</i>	+.83	+.46	-.28	-.15	1.00
<i>Sc</i>	+.72	+.08	1.37	+.20	.71
<i>AP</i>	+.85	-.14	1.18	-.12	.79
<i>AF</i>	+.77	-.33	-.24	-.41	.94
<i>Sp</i>	+.45	-.20	-.18	+.43	.46
<i>E</i>	+.76	-.20	-.22	+.17	.70

as could be expected. In view of the many things which influence progress in addition to ability, the data may be claimed to be quite consistent with a concept of general ability to learn. They do not prove its existence any more than a failure to find a general factor¹ would prove that there is no general ability to learn. *Without* a general ability, circumstances could combine to produce a general factor; *with* a general ability they could combine to obliterate the evidence of its existence.

For these data, it is thought that the loadings shown in Table 3 are more meaningful than those required for Thurstone's simple structure (5). However, for those who do not share this opinion (2), loadings after three rotations, following the Peters and Van Voorhis algebraic procedure (4), are shown in Table 4. As may be seen in Column 1, Factor 1 remains a general factor in spite of the modifications in it which were required by the transformations of Factors 2 to 4.

E. THE PATTERN OF 4-5 GAINS COMPARED WITH THE PATTERN OF PRIOR GAINS

The improvements in score made during the fourth and fifth grade experiences are referred to as 4-5 gains. The initial scores used in obtaining them

¹Factor 1 in Table 3 is Burt's General Factor (1).

TABLE 4
LOADINGS AFTER ROTATION

Area	Factor 1	Factor 2	Factor 3	Factor 4	h^2
R	.39	.09	.70	.08	.66
So	.52	.81	.16	.18	.99
L	.64	.00	.76	.15	1.00
Sc	.28	.37	.61	.35	.70
AP	.65	.22	.53	.20	.79
AF	.95	.01	.15	.00	.92
Sp	.28	.00	.05	.62	.46
E	.64	.12	.15	.50	.70

are, it is maintained, measures of gain up to the beginning of the fourth grade. They are referred to as 0-3 gains. They are the gains made from zero knowledge and ability during the pre-school years and the three years in Grades 1 through 3. To what extent does the pattern of 4-5 gains resemble the pattern of 0-3 gains?

The 0-3 gains with the exception of arithmetic fundamentals were more general than the 4-5 gains. This may be seen in Table 5, (a) in the larger

TABLE 5
COMPARISON OF FACTOR LOADINGS AT THE 4-5 AND PRECEDING LEVELS

Area	Factor 1		Factor 2		Factor 3		Factor 4	
	0-3	4-5	0-3	4-5	0-3	4-5	0-3	4-5
R	.96	.63	.19	-.18	.15	.46	-.10	-.07
So	.91	.84	.34	.52	-.06	-.09	-.07	-.15
L	.94	.83	.33	.46	-.09	-.28	-.15	-.20
Sc	.96	.72	.26	.08	-.01	.37	-.12	-.41
AP	.89	.85	-.33	-.14	-.24	.18	-.24	-.18
AF	.71	.77	-.43	-.33	-.42	-.24	-.44	-.43
Sp	.78	.45	-.15	-.20	.44	-.18	-.22	.17
E	.83	.76	-.20	-.20	.22	-.22	-.22	.17

loadings with the general factor (Factor 1), (b) in the larger communalities, and (c) in the fact that the 0-3 intercorrelations are described with three factors where four were required for the 4-5 intercorrelations. There was no doubt about three factors being sufficient. All 36 of the final residuals were within .033 of being equal to zero.

Factor 2 loadings show similar though not identical patterns. The factor is the same in showing the arithmetic and language skills to be grouped in opposition to the information subjects, social studies (history and geography), literature, and science. The factor is different in that reading was associated with the skill subjects at the 4-5 level but was associated with the information subjects at the 0-3 level.

The signs must be both plus and minus for Burt factors subsequent to the

first. Which are plus and which are minus is immaterial. The facts could be presented as satisfactorily with all signs changed. What is revealed is this. If each pupil's gains were portrayed graphically as a profile around his own average gain, it would be seen that the information gains would *tend* to be all crests, or all troughs. That is, if one were high (or low) in the profile, the others would tend to be high (or low). The same could be said of the skill gains. Also, it would be seen that for the profiles in which the crests were provided by the information subjects the troughs would be provided by the skill subjects and vice versa. In other words, on the basis of plus and minus loadings for Factor 2, one should not infer anything as to the likelihood of an area being high or low in the profile. The signs tell rather which ones tend to be together on the same side of the base line and which opposite.

What is the meaning of the difference with regard to reading? The first thing to note is that no one column tells the whole story about reading (or any subject). It would be a mistake to interpret the Factor 2 loading for reading as indicating a negative relationship between progress in ability to read and the amount of information possessed in the social studies, literature, and science areas. Such an inference would be contrary to the correlation shown in Table 2 and to the loadings with Factor 1. What it means is that the reading gains are less closely associated with general progress during the 4-5 years than they were during the 0-3 period. This could be due to the acquisition by the less able pupils during the 4-5 period of a reading skill which was acquired in the earlier period by the more able. It might be impossible for the ablest pupils to continue to improve in ability to read at the rate at which they progressed during the earlier period. At any rate, 4-5 progress in social studies (history, geography) and literature was more closely related to 0-3 gains in reading than to the 4-5 gains in reading. Progress in these information areas during the fourth and fifth grades seemed more a function of ability to read at the beginning of that period than of the gains in reading during the period.

In this difference between the 0-3 and 4-5 patterns, science gains (and to a still lower degree arithmetic problems) were associated with reading gains in being less related to general progress. Progress in arithmetic fundamentals at the 4-5 level was *more* highly correlated with general progress than it was at the 0-3 level. Thus it appears that gains in a subject-matter area may vary from level to level as to their general significance. Progress in reading, science, and arithmetic fundamentals at the 0-3 level showed no correlation with progress at the 4-5 level. The corrected r 's were .00, -.03,

and $-.03$ in these respective areas. On the other hand, for social studies, literature, arithmetic problems, spelling, and English the r 's were $.72$, $.42$, $.26$, $.32$, and $.60$.²

TABLE 6
PROGRESS IN READING AS RELATED TO 4-5 GAINS IN SOCIAL STUDIES AND LITERATURE

	Reading 0-3 Gains		Reading 4-5 Gains	
	Uncorrected r 's	Corrected r 's	Uncorrected r 's	Corrected r 's
Social Studies 4-5 Gains	.48	.64	.19	.37
Literature 4-5 Gains	.32	.50	.15	.35

This difference in relationship was so influential as to have determined the nature of the third 4-5 factor. It was only after Factor 3 accounted for this difference that the analysis revealed another similarity between the 0-3 pattern and the 4-5 pattern. The 0-3 Factor 3 and the 4-5 Factor 4 both show an association between the gains in the two parts of arithmetic, and also between the gains in spelling and the remaining parts of language usage included under the head of English. Most prominently they agree in showing opposition between arithmetic fundamentals and spelling.

When it is kept in mind that the 0-3 gains cover both a longer period and an earlier one, perhaps the patterns of intercorrelation are as similar as they should be expected to be. If they are, what ground is there for mistrusting end scores as measures of learning? For instance, for the data being described, do not (a) the initial scores, (b) the differences between initial and final scores, and (c) the final scores all have the same kind of meaning and validity? The initial score is a measure of learning up to the beginning of Grade 4. The difference between final and initial scores is a measure of learning during the fourth and fifth grades, and the final score is a measure of learning up to the end of Grade 5. *And in general, do not the conclusions about ability to learn drawn over the years from end scores have just as much validity as if they had been drawn from final minus initial scores?*

F. PREDICTABILITY OF GENERAL PROGRESS AT THE 4-5 LEVEL.

What implications are there in the data for the practice of those who could use in advance some index of the 4-5 gains in score? There are no

²These r 's were based upon two uncorrected r 's, the 0-3 Form A with the 4-5 Form B and the 0-3 Form B with the 4-5 Form A . They are thus not diminished by the negative correlation of errors which is ordinarily involved in the correlation of initial scores with gains.

direct implications concerning the use of intelligence test scores for this purpose. Intelligence test scores were not obtained. It was thought that a request for more time than was needed for the two-form achievement-testing program would have been excessive. It was also thought that if a general learning factor were discovered, substantial correlation between the general factor and intelligence test scores could be assumed, because the high com-munality between intelligence and achievement scores (3) is accepted without question. The chief question has been whether there is a general learning factor.

It can be reported that the general learning factor (Factor 1) described in Table 3 is correlated $-.52$ with chronological age. Or, to phrase it more precisely, when chronological age and an educational quotient based on initial and final scores combined were included in the matrix along with the eight measures of learning, the Factor 1 correlation was $-.52$ with chronological age and was $.72$ with educational quotient. The average uncorrected r of chronological age with the separate gains was $-.24$ and it was $-.25$ with the initial scores from which the gains were measured.

When an unweighted average of the 4-5 gains is used as the measure of general progress in the fourth and fifth grades, the possibilities for predicting it from age and achievement test data are shown in Table 7. The r 's are not high but every one of the 11 measures is an index of subsequent general improvement. The science measure alone was a better intelligence test than the whole battery and age as they are combined in FQ . Arithmetic fundamentals was the poorest index.

TABLE 7
THE POSSIBILITIES FOR THE PREDICTION OF GENERAL PROGRESS IN GRADES 4 AND 5 FROM MEASURES OF PRIOR PROGRESS

	Uncorrected r * with criterion	Corrected r with criterion
R	.38	.61
S_0	.34	.52
L	.32	.49
S_c	.45	.71
AP	.31	.47
AF	.15	.22
Sp	.29	.39
E	.37	.55
EA	.36	.50
CA	$-.31$	$-.40$
FQ	.41	.56

*The r 's in this column except for CA are averages of two coefficients Form A with Form B criterion and Form B with Form A criterion. The CA r is the average of the r between CA and Form A criterion and the r between CA and Form B criterion.

G. SUMMARY AND CONCLUSION

"Scholastic ability is general. Get a good measure of a pupil's ability to learn, and you may expect him to measure up to it in all the major subjects." This line of thinking has been used very commonly and confidently at all scholastic levels. Only in recent years has a question been raised as to the validity of a concept of a general ability to learn. The questioning is desirable and profitable (6). The present article and the earlier one referred to above were not designed as a protest against critical examination. Protest is directed solely to the use of poor and inappropriate measures. Considerable space was therefore given above to the discussion of factors which might diminish or even prevent positive intercorrelation between measures of school learning. Data were secured and selected in which the prospects were thought to be fairly good for a general ability to manifest itself, if there is one. It was pointed out that a generalness of progress does not constitute *proof* of a general ability to learn, any more than lack of agreement constitutes disproof.

The intercorrelations between the measured gains were all positive and substantial. The Burt first or general factor remained a general factor when the three bipolar factors were rotated to have zero to positive loadings. The gains in score were not as general as were the initial scores, themselves measures of what had been learned up to the time of the initial test. Less generality at an upper level is not inconsistent with the idea of a general ability to learn, for a general ability is not inconsistent with an increasing specialization.

The pattern of intercorrelation at the two levels, during the fourth and fifth grades and before, was similar. This was in spite of the facts that the periods differed in length, that one was partly pre-school, and that the measures were end-tests at one level and differences between initial and final scores at the other level. Such differences as appeared are interpreted as examples of ways in which the teaching situation may prevent a general ability to learn from manifesting its influence. It seems, for instance, that to a certain extent the fourth and fifth grades were a period for the less able pupils to catch up in their ability to read while the more able pupils used the ability they had already acquired. At any rate, progress at the fourth and fifth grade levels in three areas, reading, science, and the fundamental operations of arithmetic, was uncorrelated with progress in these areas prior to the fourth grade. With these three exceptions, the tendency was for greater progress in the fourth and fifth grades to be made by those who had made the greater prior progress.

On the whole, the data presented here, and in the earlier article, are fairly consistent with a concept of a general ability to learn and with the identification of it with the "general" intelligence test.

REFERENCES

1. BURT, C. *The Factors of the Mind*. New York: Macmillan, 1941.
2. CATTELL, R. B. *Factor Analysis*. New York: Harper, 1952.
3. KELLEY, T. L. *Interpretation of Educational Measurements*. Yonkers-on-Hudson: World Book, 1927.
4. PETERS, C. C., & VAN VOORHIS, W. R. *Statistical Procedures and Their Mathematical Bases*. New York: McGraw-Hill, 1940.
5. THURSTONE, L. L. *Multiple Factor Analysis*. Chicago: Univ. Chicago Press, 1947.
6. TILTON, J. W. *An Educational Psychology of Learning*. New York: Macmillan, 1951.
7. ———. Intelligence test scores as indicative of ability to learn. *Educ. & Psychol. Meas.*, 1949, 9, 291-296.

Department of Education
Yale University
28 Hillhouse Avenue
New Haven, Connecticut