Genetic Experiments with Animal Learning: A Critical Review

DOUGLAS WAHLSTEN

Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada

The basic patterns of inheritance of learning ability in animals have been delineated. Summaries of strain differences in learning rate, responses to selective breeding for learning, heritabilities of learning phenotypes, and heterosis and overdominance are presented. In addition, the patterns of inheritance are shown to vary with the early environment.

The causes of genetic differences have received much attention, but much of the research is inconclusive. Both general learning ability and task-specific abilities are important, but their relative importance is not known for most learning tasks. Strain differences have been found to vary widely in response to variations in stimulus parameters, motivational levels, temporal spacing of trials, and pharmacological manipulations. However, in only a few cases have strain differences in learning actually been shown to be attributable to differences in sensory capacities, motivation, memory or activity levels. The physiological bases for differences are totally unknown. The pathways of gene action on learning also await discovery.

Although some researchers have claimed to study the adaptive value of learning, their exclusive utilization of laboratory populations precludes meaningful interpretation of their results.

Several methodological shortcomings of various experiments are considered, and important areas for future research are suggested.

Learning is a phenotype which has engaged the interests of numerous researchers seeking genetic bases for behavioral differences. In fact, much of the earliest research identifiable as behavior genetics dealt with some aspect of learning in animals (Bagg, 1916; Yerkes, 1916; Tolman, 1924). Ensuing experimentation was performed primarily by psychologists using genetically ill-defined populations. The rather recent appearance of standardized inbred

1Preparation of this paper was supported in part by Grant APA-398 from the National Research Council of Canada.
2The review is not exhaustive, since only directly pertinent studies are presented. However, a supplementary bibliography is included at the end of the paper which contains other relevant literature. A more complete review is available from the author upon request.
mouse strains with widespread availability has led to renewed interest in the genetic analysis of learning, as well as other phenotypes. Several sophisticated quantitative genetic tools are now readily available for the study of learning. Examples of the application to learning research of selective breeding, the classical cross, the diallel cross, sib analysis, parent-offspring regression, and single-gene analysis have appeared recently.

A central motive for compiling the present review is the author’s opinion that the increase in genetic sophistication has not been paralleled by a similar growth in the sophistication of measures of learning. In many studies it appears that learning was selected as a phenotype of convenience and general interest.

Similarly, the questions about learning investigated have tended to be simplistic and of little interest to those concerned with the nature of the learning process itself. Many recent studies have raised issues that were presented in the earliest research and therefore have contributed little to progress in the area. This is an unfortunate situation in view of the potential power of genetic techniques to answer important questions about learning. It is hoped that the organization of the present review around major questions in the area, instead of around techniques or species, will clarify some of the issues and indicate promising directions for future research.

THE NULL HYPOTHESIS

Logically, although not chronologically, the first issue to be raised is whether genes affect learning at all. To the student of animal behavior in 1971, it seems a little unbelievable that informed scientists ever seriously questioned the involvement of genotype in the learning process, given that genetic effects upon physical and chemical characteristics were so widely known. Nevertheless, this was a very lively issue until quite recently, and it spawned numerous experiments which purported to demonstrate that animals known to have different genotypes also had different scores on a particular learning task. Even today such experiments continue to be performed and subsequently are considered worthy of promulgation.

Strain Comparisons

The first step in examining genetic differences in learning is, of course, to obtain some animals which are known by other criteria to possess different genotypes. This is most easily done by procuring standardized strains which have been inbred for at least 20 generations, using brother-by-sister matings to ensure that less than 2% of the loci are likely to be unfixed. Similar comparisons of noninbred animals are also pertinent, although the various haphazard breeding schemes and diverse origins of the parent populations used
to maintain the lines preclude the possibility of guaranteeing samples with uniform gene frequencies in successive generations or even different shipments from the same supplier and thereby prevent finer analyses of observed differences.

Several strain comparisons of performance on learning tasks are summarized in Table 1. All of the cited studies reported significant between-strain differences; most differences were highly significant (i.e., \( p < .001 \)). The results were obtained with a wide range of training procedures using both appetitive and aversive motivation. Although a relatively narrow range of laboratory animal populations has been studied, it is clear that significant genetic variation in learning is to be expected as the rule rather than as the exception. The rare experiments leading to negative results generally involved only two strains and therefore possessed minimal power (see footnote 2).

**Artificial Selection**

In a heterogeneous population composed of very many genotypes, virtually one per individual, artificial selection for high and low learning scores is a very strong test of genetic involvement in learning.

All of the early selection studies used rats in the exceedingly complex mazes in vogue at the time, and they all had two purposes: to produce lines of rats with high and low error scores and to fix these lines for the loci relevant to learning by the process of inbreeding. The first goal was to show that genes affected learning ability, and the second was presumably to allow subsequent analyses of the genetic mechanisms involved. Tolman (1924) selected for high and low scores based upon a “rough pooling” of errors, running time, and number of perfect runs in a complex maze for two generations. Although the two selected lines were significantly different in both the F1 and F2 generations, the intrasubject reliability of the maze test was so close to zero that Tolman abandoned his effort. Determined to avoid some of Tolman’s problems, Tryon (1929) selected for high and low error scores on a 17-unit maze of known, high reliability (.95); he also reduced inbreeding by using only 50% full-sib matings. He later added high fertility, good health, and coat color to the selection criteria. His results, which are widely known among psychologists, showed clear divergence of the two lines, such that very little overlap existed by the last generation of selection, F22 (Tryon, 1940). The maze bright and dull strains, termed S1 and S3, have been maintained since then by random breeding and are still available today. A very similar selection study was conducted by Heron (1935) using an automatic Minnesota 12-unit maze, and very similar results were obtained. Whereas the parent population averaged about 85 errors on trials 3 through 17, by F16 there was almost no overlap, the mean errors being 46.9 for the brights and 116.0 for the dulls (Heron, 1941). For some reason the brights were very superior even on the first trial, while the rate of error reduction was about the
### TABLE 1

Summary of Several Studies of Strain Differences in Learning Rate

<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Motivation</th>
<th>Strains</th>
<th>( \omega^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Royce and Covington (1960)</td>
<td>Shuttle</td>
<td>Shock</td>
<td>9</td>
<td>.18(^b)</td>
</tr>
<tr>
<td>Lindzey and Winston (1962)</td>
<td>T maze</td>
<td>Hunger</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Meier and Foshee (1963)</td>
<td>Maze</td>
<td>Water escape</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Collins (1964)</td>
<td>Shuttle</td>
<td>Shock</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Wimer and Weller (1965)</td>
<td>T Maze</td>
<td>Water escape</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Schlesinger and Wimer (1967)</td>
<td>Jump-up</td>
<td>Shock</td>
<td>7</td>
<td>.45</td>
</tr>
<tr>
<td>Henderson (1968a)</td>
<td>CER</td>
<td>Shock</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Bovet et al. (1968)</td>
<td>Shuttle</td>
<td>Shock</td>
<td>9</td>
<td>.95</td>
</tr>
<tr>
<td>Bovet et al. (1969)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bovet-Nitti (1969)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carran (1969)</td>
<td>T maze</td>
<td>Thirst</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Fuller (1970)</td>
<td>Sidman shuttle</td>
<td>Shock</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Stasik (1970)</td>
<td>T maze</td>
<td>Shock</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Wahlsten (1971)</td>
<td>Jump-out</td>
<td>Shock</td>
<td>4</td>
<td>.18</td>
</tr>
<tr>
<td>Wahlsten (1971)</td>
<td>One-way</td>
<td>Shock</td>
<td>4</td>
<td>.34</td>
</tr>
<tr>
<td>Wahlsten (1971)</td>
<td>Optional</td>
<td>Shock</td>
<td>4</td>
<td>.11</td>
</tr>
<tr>
<td>Wahlsten (1972)</td>
<td>Jump-out</td>
<td>Shock</td>
<td>7</td>
<td>.37</td>
</tr>
<tr>
<td>Wahlsten (1972)</td>
<td>One-way</td>
<td>Shock</td>
<td>7</td>
<td>.42</td>
</tr>
<tr>
<td><strong>Rats</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sawrey and Long (1962)</td>
<td>Ulcer formation</td>
<td>Shock</td>
<td>4(^c)</td>
<td></td>
</tr>
<tr>
<td>Harrington (1966)</td>
<td>Elevated maze</td>
<td>Hunger</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Wilcock and Broadhurst (1967)</td>
<td>Shuttle</td>
<td>Shock</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Harrington (1968)</td>
<td>Hebb-Williams maze</td>
<td>Hunger</td>
<td>11</td>
<td>.36</td>
</tr>
<tr>
<td>Anisman and Waller (1972)</td>
<td>One-way</td>
<td>Shock</td>
<td>5(^c)</td>
<td></td>
</tr>
<tr>
<td>Anisman and Waller (1972)</td>
<td>Shuttle</td>
<td>Shock</td>
<td>5(^c)</td>
<td></td>
</tr>
<tr>
<td><strong>Dogs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fuller (1955)</td>
<td>Leash control</td>
<td>?</td>
<td>5(^c)</td>
<td>.31</td>
</tr>
<tr>
<td>Freedman (1958)</td>
<td>Inhibition</td>
<td>Swat and scold</td>
<td>4(^c)</td>
<td></td>
</tr>
<tr>
<td>Elliot and Scott (1965)</td>
<td>T maze</td>
<td>Hunger</td>
<td>5(^c)</td>
<td></td>
</tr>
<tr>
<td>Scott and Fuller (1965)</td>
<td>Several</td>
<td>Several</td>
<td>5(^c)</td>
<td>.27</td>
</tr>
<tr>
<td><strong>Chickens</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altevogt (1951)</td>
<td>Visual discr.</td>
<td>Hunger</td>
<td>6(^c)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Only studies reporting data for four or more strains are included.
\(^b\)Calculated only for five strains which learned within 700 trials.
\(^c\)Noninbred strains.

same for the two strains. Finally, Thompson (1954) selected for “intelligence” by administering 24 different problems on the Hebb-Williams maze; he also used full-sib matings exclusively until F\(_6\). The error scores of the high and low lines diverged significantly, but by F\(_6\) so many matings were infertile that inbreeding had to be abandoned.
Since these early efforts, psychologists have become aware that the two goals of selection, high- and low-scoring genotypes and genetic fixation by inbreeding, are diametrically opposed. Selection operates on genetic variance, which is progressively reduced by inbreeding. This is not to say that no response to selection will occur if inbreeding is practiced, but the rate of divergence and the asymptotic separation of the two selected lines will certainly be reduced. In addition, inbreeding can lead to sterility of many matings and even loss of the selected lines altogether.

Realizing this, Bignami (1965) selected for high and low scores on avoidance in a shuttlebox without using any full-sib pairs; he also tried, but lost, a line selected with concurrent inbreeding. A large response to selection was observed in the very first selected generation, and even larger separation of lines was obtained by the fifth generation. The parent population averaged 104.9 avoidances in 250 trials, while by $F_5$ the high line had a mean of 170.6 avoidances compared to 50.9 for the low line. No difficulties with sterility were reported for either of the lines. Bovet et al. (1969) also obtained a rapid response to selection for shuttle avoidance learning in mice, although they reported only a line selected for high scores. Finally, Schaefer (1968), believing that response duration was a determinant of intelligence, selected for the time required to perform 100 lever presses on an FR10 schedule for food reward in mice. He reported two generations of selection for long and short times with no sib matings. In both generations there was a significant difference ($p < .01$) between the two lines.

It is evident that success in selectively breeding for high and low learning rates in laboratory rats and mice is commonplace. Taken together with the numerous strain comparisons mentioned above as well as more sophisticated genetic experiments to be presented below, these results allow the null hypothesis that genotype does not affect learning to be firmly rejected for the populations studied.

RELATIVE MAGNITUDE OF GENETIC VARIATION

Once the statistical significance of learning differences between animals of the various genotypes has been firmly established, the question arises concerning the relative importance of genetic variation as a source of variation in learning ability. If large numbers of subjects from numerous strains must be tested to establish the validity of the phenomenon, then the importance of genetic variation is questionable. On the other hand, if a substantial portion of the total variation in learning scores within a population of animals can be traced to genetic origins, then students of learning must give serious attention to the genetic structure of their experimental populations.

The question of relative importance can be stated quite simply: What proportion of the total variance in a learning phenotype in a population can be attributed to genetic differences among individuals? In the case of strain
comparisons with a one-way analysis of variance design, this question can be answered by calculating the strength of effect ($\omega^2$). In experiments involving breeding, it is customary to posit a linear model for genetic effects and then partition variances appropriately. If an individual's score or phenotype ($P$) is partitioned into components of genetic ($G$) and environmental ($E$) origin, and if $G$ and $E$ do not interact, then $P = G + E$, and the variances are such that $V_P = V_G + V_E$ (see Roberts, 1967a, for a more complete presentation). The relative contribution of genetic differences is $V_G/V_P$; this ratio is sometimes termed the coefficient of genetic determination ($C_G D$). A valid measure of this coefficient necessitates that the effects attributable to $G$ and $E$ be clearly distinguishable. For a multitude of reasons, direct measures of this ratio are not easily obtained. However, a related measure, heritability, has similar properties and can be estimated accurately. Heritability ($h^2$) is the ratio of additive genetic variance to total phenotypic variance. Additive variance ($V_A$) is a manifestation of the average values of genes at each relevant locus as opposed to nonadditive effects such as dominance ($D$) and interaction between loci ($I$, epistasis). Since $V_G = V_A + V_D + V_I$, additive variance is always less than or equal to total genetic variance, and heritability is always a conservative estimate of the relative contribution of all genetic differences.

**Strength of Effect**

The coefficient $\omega^2$ estimates the proportion of total variance in an experiment which can be attributed to differences between strains. When highly inbred strains are employed, between-strain variation should reflect primarily genetic variation, while within-strain variation should represent differences in postfertilization environment as well as error in measuring the behavior itself.

Several estimates of $\omega^2$ for strain comparisons are presented in Table 1. The values were derived from the $F$ ratio for between-strains differences and the degrees of freedom between ($df_b$) and within ($df_w$) strains. It can be shown that the expression for estimating $\omega^2$ given by Hays (1963, p. 382) reduces to

$$\text{est. } \omega^2 = \frac{F-1}{F + [(df_w + 1)/df_b]}$$

for a one-way design with equal numbers of subjects per cell. It should be noted that only two reports (Scott and Fuller, 1965; Wahlsten, 1971) actually presented values for $\omega^2$. The remainder were derived by the present author.

The wide range of estimated $\omega^2$ values indicates that no simple statement can be made. It should be noted, however, that many values greater than 30% were obtained, which signifies a very substantial effect as judged by results from other areas of behavioral research.
Coefficient of Genetic Determination

Oliverio, Castellano, and Messeri (1971) have presented the only calculations of C.G.D. for a learning phenotype. They found C.G.D. for percent correct in 500 trials of shuttle avoidance learning to be .64 for a cross of inbred mouse strains SEC/1ReJ and C57BL/6J and .84 for the cross of DBA/2J and C57BL/6J. Corresponding values for total errors in 15 trials of a Lashley III maze were .50 (S X C) and .39 (D X C).

Heritability

Of the several methods available for calculating heritability ($h^2$), realized response to selection for learning appears to be the most efficient (Hill, 1971). It is unfortunate that the various selection studies mentioned above were improperly designed to allow estimation of realized heritability. Some of these difficulties are evident in Table 2, which lists several pertinent aspects of the experiments. A proper selection experiment by DeFries and Hegmann (1970) involving open field activity in mice is included in the table for purposes of comparison.

No researcher can obtain today a population known to have the same genetic properties as any of those previous ones, because the breeding schemes employed by most animal suppliers are generally haphazard and are certainly not uniform for different suppliers of the same outbred strains. Also, in all studies, except those of Schaefer (1968) and Bovet et al. (1969), the selection criterion was a composite of the learning score of primary interest and some other trait such as running time or fertility. This means that the response to selection no longer has a simple relation (i.e., heritability) to cumulated selection differential; it is instead dependent upon the heritability of the composite and the genetic correlation between learning and the other components of the selection criterion. Finally, none of the experiments utilized an adequate unselected control line, which is quite important for minimizing the effects of environmental changes from one selection generation to the next and for detecting an asymmetrical response (DeFries, 1967). These several shortcomings may be contrasted to the DeFries and Hegmann experiment, in which repeatability was assured by the adoption of a cross between genetically fixed inbred strains, a single response measure served as the selection criterion, inbreeding was minimized, and replicated control and selected lines were included.

Other methods for estimating $h^2$ (see Roberts, 1967a; Falconer, 1960) have been employed with greater success. These studies are summarized in Table 3 together with estimates from two selective breeding studies. It is interesting that heritability measures show a smaller range (.2 to .5) than values of $\omega^2$ in Table 1 (.1 to .95). It is also interesting that four experiments with shuttle avoidance learning using four highly dissimilar populations found $h^2$ values of about .5.
<table>
<thead>
<tr>
<th>Parent population</th>
<th>Number of lines</th>
<th>Selection criterion</th>
<th>Selected generations</th>
<th>Breeding scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source</strong></td>
<td><strong>Number</strong></td>
<td><strong>High</strong></td>
<td><strong>Low</strong></td>
<td><strong>Control</strong></td>
</tr>
<tr>
<td>Tolman (1924)</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Albino rats,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berkeley anatomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dept.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tryon (1929)</td>
<td>142</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Rats from many</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unrelated litters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heron (1935)</td>
<td>44</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mixture of hooded</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>and albino rats</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thompson (1954)</td>
<td>??</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Albino and hooded</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rats</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bignami (1965)</td>
<td>62</td>
<td>1</td>
<td>1</td>
<td>1*</td>
</tr>
<tr>
<td>Wistar albino rats</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schaefer (1968)</td>
<td>64</td>
<td>1</td>
<td>1</td>
<td>1*</td>
</tr>
<tr>
<td>Swiss-Webster mice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bovet <em>et al.</em> (1969)</td>
<td>32</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Swiss-Webster mice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeFries and</td>
<td>497</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Hegmann (1970)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>F_3</strong> of BALB/cJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>× C57BL/6J</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

*It is difficult to imagine how 44 rats in the parent population could provide 15 males and 30 females per line or even for total matings. Heron's comments (1935, p. 81) suggest that the 30 matings included both the bright and dull lines.*

*b“Control” line purchased from supplier of original parent population at a later time.*
### TABLE 3
Several Estimates of Heritability of Learning in Animals

<table>
<thead>
<tr>
<th>Study</th>
<th>Parent population</th>
<th>Task</th>
<th>Genetic method</th>
<th>$h^2 \pm SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willham et al. (1963)</td>
<td>Duroc and Hampshire swine</td>
<td>Shuttle avoidance</td>
<td>Sib analysis</td>
<td>.45 ± .12</td>
</tr>
<tr>
<td>Bignami (1965)</td>
<td>Wistar rats</td>
<td>Shuttle avoidance</td>
<td>Artificial selection</td>
<td>.56^a ± .15</td>
</tr>
<tr>
<td>Schaefer (1968)</td>
<td>Swiss-Webster mice</td>
<td>FR10 lever pressing for food</td>
<td>Artificial selection</td>
<td>.34 ± .12^b</td>
</tr>
<tr>
<td>Henderson (1968a)</td>
<td>Four inbred mouse strains</td>
<td>CER with shock US</td>
<td>Diallel cross</td>
<td>.21 ?</td>
</tr>
<tr>
<td>Tyler and McClearn (1970)</td>
<td>HS/Ibg mice</td>
<td>Straight-alley running for food</td>
<td>Regression</td>
<td>.30^c ± .10</td>
</tr>
<tr>
<td>Oliverio (1971)</td>
<td>CD1 mice</td>
<td>Shuttle avoidance</td>
<td>Sib analysis, regression</td>
<td>.50 ± .10</td>
</tr>
<tr>
<td>Oliverio et al. (1971)</td>
<td>Three inbred mouse strains</td>
<td>Shuttle avoidance</td>
<td>Sib analysis, regression</td>
<td>.48 ± .08</td>
</tr>
<tr>
<td>Oliverio et al. (1971)</td>
<td>Three inbred mouse strains</td>
<td>Lashley III maze for food</td>
<td>Sib analysis, regression</td>
<td>.40 ± .06</td>
</tr>
</tbody>
</table>

^a Calculated by the present author from regression of cumulative response on cumulative selection differential for high line (RHA). Low line (RLA) showed a large response in the first generation of selection but great variability thereafter; the regression coefficient for RLA was calculated to be +.08.

^b Standard error derived from limits of 95% confidence interval given by Schaefer (1968).

^c Heritability of slope of line of best fit to latency decrease across five days of training.

Although the proper interpretation of these measures of $\omega^2$, $C.G.D.$, and $h^2$ is not readily apparent, some limitations on their generality are obvious. The inherent genetic variation of a population influences greatly the results, since reduction of $V_G$ through inbreeding or of $V_A$ through selection would lead to the observation of low $h^2$. Similarly, environmental attributes can influence the $V_E$ component. Intuitively, rearing under uniform conditions is expected to yield the largest possible proportion of genetic variance, because $V_E$ should be small. However, recent evidence reported by Henderson (1970) clearly demonstrates that the typical restrictive laboratory environment may actually suppress the manifestations of genetic variation and thereby yield a lower heritability score than would otherwise be obtained if the
animals were raised in an enriched environment. Thus, the magnitude of the heritability coefficient is affected by the environment of the subjects as well as their actual genetic variation and, as a result, cannot be relied upon to be invariant in other worlds.

Another factor must be the reliability of the learning measure itself. If the environmental component, "E," is partitioned into $E$ due to pretesting environment and $e$ from noise in the measuring instrument, it follows that $V_P = V_G + V_E + V_e$. $V_e$ will be small for tests with high test-retest reliability ($r_{tt}$) or when many repeated measures on the same animals are administered. The data presented by Bovet, Bovet-Nitti, and Oliverio (1969, p. 140) show that individual scores in shuttle avoidance are very stable from day to day when 100 trials are administered; in turn they find large strain differences ($\omega^2 = .95$, Table 1). On the other hand, experiments which examined relatively short learning sequences of only a few trials (Henderson, 1968a; Wahlsten, 1971) reported lower values of $h^2$ (.2) and $\omega^2$ (.1). Estimation of $r_{tt}$ will aid the interpretation of $h^2$ in the future.

The magnitude of $\omega^2$ and $h^2$ may also be influenced by the difficulty of the task employed. Wahlsten (1971) found that requiring mice to either run (one-way) or jump (jump-out) led to $\omega^2$ values of .34 and .18, respectively, but that a smaller $\omega^2$ of .11 resulted when each subject could either run or jump (optional) to escape or avoid shock (see Table 1). Other simple tasks such as CER conditioning (Henderson, 1968a) and straight-alley running (Tyler and McClearn, 1970) show low heritabilities (.2 to .3), while the more difficult shuttle avoidance yields C.G.D. of over .6 and $h^2$ of about .5. Thus, genotypes which are all sufficient for learning simple tasks may not be equally effective when the demands for processing information are increased. Since the above studies provide only indirect evidence, this idea should be subjected to direct testing in the future. It will be necessary to devise a battery of tests in which only task difficulty is varied without changing the source of motivation, the relevant sensory modality, or the motor response requirements.

Another important aspect of heritability is its relation to fitness and the adaptive value of learning ability. This topic will be discussed in another section of the paper.

GENETIC CORRELATES OF LEARNING

Observation of large genetic variation in learning rates leads directly to questions about the causal bases for these differences, as well as their generality to other kinds of learning. It is worthwhile to determine precisely what mechanisms or components of the learning process are modified in different gentotypes and thereby yield the observed phenotypic differences. If there exists a finite set of mechanisms that results in overt learning, are all of
these mechanisms affected by genetic variation, or are certain components of
the learning process more likely to be changed than others?

Whenever a complex behavior such as learning is the object of study,
many genes are expected to be involved in differences between genotypes.
Although no one gene may be individually identifiable, it is possible to study
relations between polygenic traits with the methods of quantitative genetics.
While pleiotropic gene action at any one locus may not be demonstrable, the
genetic correlation coefficient measures something analogous to pleiotropy. In
order to accomplish this, it is necessary to perform a genetic experiment by
crossing individuals that differ in genotype. If an experiment is correctly
designed and executed, it is possible to partition the correlation between two
phenotypes ($r_p$) into components attributable to genetic similarities ($r_g$) and
environmental actions ($r_e$). Actually the more common practice is to partition
between additive genetic similarities ($r_A$) and everything else (“$r_E$”). Falconer
(1960) showed that the appropriate relation is $r_p = h_x h_y r_A + e_x e_y r_E$, where
$x$ and $y$ are the phenotypes being compared, $h$ is the square root of
heritability ($h^2$), and $e = \sqrt{1 - h^2}$. Several things are apparent from this
relation. The correlation of phenotypes may be the result of covariation in
either genotype, environment, or both. No conclusive statement can be made
a priori; the actual magnitudes of $r_A$ and $r_E$ must be estimated with a genetic
experiment. Furthermore, the contribution of genetic covariation to pheno-
typic similarity may be small if heritability of either phenotype is small. The
value $r_A$ is commonly interpreted as a measure of the proportion of genes
which are intersecting subsets of the sets of genes affecting each trait. If $r_A$
is high, approaching 1.0, then the two traits are probably controlled by the very
same physiological mechanisms, whereas low values of $r_A$ indicate that the
two traits are controlled by independent sets of genes and mechanisms.

Several methods have been employed to study the genetic correlates of
learning. Since they are not equally useful, it is pertinent to discuss briefly
their limitations at the outset.

The simplest design applicable to this question entails the measurement
of many other characteristics of strains of animals that are already known to
differ on at least one learning task. More elegant experiments subject the
strains to different experimental manipulations in order to determine whether
all strains are affected equally or whether the original differences in learning
are to be found under other conditions. However, the nature of gene fixation
during inbreeding leads one to believe that the study of inbred strains alone
can never reliably detect the causes of learning differences, regardless of the
outcome of an experiment. Briefly stated, it is utterly impossible to determine
whether two distinct behaviors observed in a single genotype (i.e., an inbred
strain) are controlled by identical, overlapping, or entirely independent sets of
genes by the sole method of statistical comparisons of several strains. Even if
a significant and substantial correlation between two phenotypes occurred, it
still could not be confidently stated that a causal genetic relation existed, for
they might be similar for reasons other than common genetic mechanisms of action. They might be manifestations of common experience, if the measures come from the same animals.

The simple operation of crossing inbred strains to obtain $F_1$, $F_2$, and backcross generations provides an abundance of information which cannot be obtained by any environmental manipulations of inbred strains alone. Paramount among these benefits is the possibility of examining correlations between several aspects of learning which were observed to covary among the parent strains. When the strains are crossed, the measures of learning or other behaviors in the $F_1$ and $F_2$ generations may continue or cease to exhibit phenotypic correlations, depending on whether they are genetically related or independent, respectively.

James (1941) seems to have been the first to employ this technique to study correlations. He observed correlations between body type and learning of leg-flexion avoidance and Pavlovian salivation training. The outcome of crossing two breeds was clear:

In the two polar types... there seems to be a definite correlation between bodily form and behavior. There is a harmonious relationship among the genetic factors for physical form, glandular conditions, and behavior. When the two polar types are bred together, however, this relation breaks up. A dog may inherit the bodily form of the basset hound, yet behave like the excitable shepherd dog under experimental conditions (p. 613).

Whereas a strain study may detect concommitants to learning differences which really are quite unrelated to learning, a proper selection study in which a learning phenotype is the only selection criterion will lead to correlated changes in other phenotypes that are related to learning through the additive action of common genes. By employing large enough populations in the selected lines, spurious correlations resulting from random sampling or genetic drift may be reduced to a very small magnitude. Correlated responses to selection become especially informative in such an experiment because the ones most closely related genetically to the learning genotype should show the most rapid response to selection, while measures that are less closely related should exhibit correspondingly smaller changes. Thus, in principle, the selection experiment can be employed to derive empirically the additive or linear genetic correlates of learning ability.

It must be mentioned that most of the above selection studies were not conducted in a manner that allowed computation of $r_A$. Parent populations and selected lines tended to have few animals (see Table 2), and control and replicated selection lines were omitted.

The most useful techniques for the study of genetic correlates entail the study of parents and offspring in a random-breeding population. They allow robust estimates of both $r_A$ and $r_E$ between phenotypes, and the accuracies of these estimates may be calculated easily.
Since the interest of most researchers centers on learning ability in the broader sense rather than on performance changes during a single training procedure, it is important to determine whether strain differences with one task are also observable with other paradigms and motives. General learning ability in animals may be analogous to the concept of intelligence (g) in humans and in this respect is a measure which should transcend the specific requirements of any one task.

Bovet et al. (1969) reported that the rank ordering of nine mouse strains on a shuttle-avoidance task was very consistent with the relative abilities of the strains in Lashley III maze learning (Spearman $r = .92$). Since the two training procedures were vastly different, the similar ordering of strains suggested that the genetic differences affected learning at a quite general level. On the other hand, Fuller (1970) tested four inbred mouse strains on either active or passive shuttle avoidance with a procedure that used no discriminative CS. Strain rank orders were completely inverted for the two procedures. Pharmacological manipulations suggested that activity or “kinetic drive” differences were more important than any differences in general learning ability. Resolution of these seemingly divergent findings has been made possible by the recent work of Oliverio et al. (1971) mentioned above. They calculated genetic correlations between shuttle avoidance learning, Lashley III maze learning, and wheel running activity. The $r_A$ between shuttle and maze learning was about $.73 \pm .12$, indicating that common abilities are required for both tasks but that unique aspects exist as well. One of these “unique aspects” for shuttle avoidance was wheel-running activity, for $r_A$ between these two was about $-.71 \pm .12$, which implies that high “kinetic drive” may interfere with discrete-trial avoidance learning. Wheel running was not related to maze learning.

One feature of the literature on strain variation in avoidance learning appeared to argue against any significant general learning ability. The problem was that some investigators observed certain strains, e.g., C3H or CBA, to learn very slowly, if at all (Bovet et al., 1968; Bovet-Nitti, 1969), while others found the same strains to be among the best learners (Stasik, 1970; Collins, 1964). Wahlsten (1971) obtained this result within one experiment; the CBA/J strain learned jump-out avoidance most quickly but was very poor at one-way avoidance. Subsequent genetic analyses (Wahlsten, 1972) demonstrated that the interaction was caused by the gene retinal degeneration ($rd$). When effects of $rd$ and albinism ($c$) were eliminated, strain ranks were similar with the two procedures.

Although the above experiments with inbred mice indicate the importance of general learning ability, research with other species has frequently revealed substantial strain-by-training procedure interactions. Harrington
WAHLSTEN

(1968) reported that certain rat strains were much better on certain problems of the Hebb-Williams maze but were inferior on other problems. Pryor and Otis (1970) found that rats of the Buffalo strain achieved criterion more quickly than Fischer rats for successive brightness discrimination in an underwater T maze but that the Fischer strain was superior at pole-displacement avoidance learning. James (1941) subjected basset, German shepherd, and saluki dogs, which were rated as lethargic, active, and very active, respectively, to restraint in a conditioned reflex stand and then to leg flexion avoidance training. The lethargic bassets submitted easily to restraint, required intense shock to elicit a leg flexion, and never performed the avoidance consistently, whereas the German shepherds struggled violently when restrained but learned to avoid very rapidly. He later trained similar groups of dogs on conditioned salivation and then on leg-flexion avoidance (James, 1953). The active dogs gave poor conditioned salivary responses but were good at avoidance, although some struggled to a degree which made reliable measurement of any learning quite impossible. The lethargic types had good salivation responses early in training, but they tended to fall asleep later; they seldom learned to avoid. However, dogs of medium activity demonstrated both good salivation and proficient avoidance. Dykman, Murphree, and Peters (1969) also observed interactions with their bold and friendly (A) and timid (E) strains of pointer dogs. When operant bar-press training for food reward was given, 31 of 34 A dogs performed at a moderate to high operant level, while 30 of 48 E dogs failed to acquire even a modest rate of bar pressing. During classical leg-shock conditioning, however, the E dogs achieved a significantly higher frequency of conditioned leg-flexion responses to a 500-Hz tone. In contrast to the skeletal motor CR measure, heart rate revealed a superior discrimination between positive and negative tones for the A dogs. Similar results were obtained for respiration rate. Thus, the measure of learning determined to a large extent which strain of dogs was judged to have superior learning ability. Strain interactions may also attenuate the generality of statements based upon group data when genetically variable dog populations are studied (see Wahlsten and Cole, 1971).

The learning abilities on diverse tasks of strains selected for learning rate on a single task are also of interest. Schaefer (1968), who selected for response duration in lever pressing, found that the mice with shorter response durations did in fact learn a T maze faster than the more persevering strain. This supported Schaefer's contention that response duration was an important determinant of intelligence.

More extensive tests have been performed with the descendants of Tryon's lines (Brights are $S_1$, Dulls are $S_3$). Certainly, the most eminent study among these was by Searle (1949), who measured each subject on numerous maze tasks and other behaviors in addition to the original Tryon maze. Appropriately enough, $S_1$ was quite superior to $S_3$ on the original Tryon maze, and it was better on a 14-unit elevated maze as well, although some
overlap existed in the latter scores. However, the S3 were superior to S1 rats in the water-escape tank, while no difference was apparent in the 16-unit and 6-unit discrimination tasks. The pattern of scores led Searle to suggest that a motivational difference existed, the S1 strain being more highly motivated by hunger and the S3 by water-escape. Rosenzweig, Krech, Bennett, and Longwell (1958) tested S1 and S3 on the Hebb-Williams, Dashiell, and Lashley III mazes using food reward and found the S1 strain to be superior on all three. Fehmi and McGaugh (1961) found that S1 learned a horizontal-vertical discrimination faster than S3, but they found no difference in black-white discrimination learning. Their result was extended when Wolfer (1963) observed that S1 exhibited fewer errors on a Lashley III maze than S3 at each of three different deprivation levels but that the two always had similar running times. In several recent studies avoidance learning has been tested as well. The S3 rats were better at avoidance learning in an ATLAS maze with visual cues (Markowitz and Sorrells, 1964) but not with spatial cues (Markowitz and Becker, 1965), while the S1 strain seemed to be superior in wheel-turn avoidance (Zerbolio et al., 1965) but inferior in jump-out avoidance (Powell and Leach, 1967).

Thus, research with the Tryon strains has confirmed the findings of the many strain comparisons in that reversals in learning rates may occur when strains are tested on tasks having many differences. The existence of such interactions makes it imperative that the degree of genetic correlation between tasks be quantified as was done by Oliverio et al. (1971). The wisdom of extending these methods to a larger number of strains and tasks in future research needs no emphasis.

Of course, learning rate is one thing, but a full-blown law of learning is quite something else. Strains could differ widely in acquisition rates on diverse tasks without necessarily invalidating learning principles. A principle can be studied only by experimental manipulation of several independent variables which are believed to influence learning and performance. Since most of the studies reviewed herein were relatively modest in their use of independent variables, it is clear that most researchers were not interested in this particular question. The more extensive experiments generally did not test anything resembling a law of learning. Hence, judgment must be suspended for lack of evidence.

Lest there be a sudden upsurge in behavior-genetic analyses of learning principles, researchers should be aware of the current state of flux in the study of learning by the more traditional methods of psychology. Seligman (1970) questioned the principle of equal associability of all stimuli and responses using any reinforcement. He suggested that the laws of learning apply only to those responses which organisms are prepared to make to certain stimuli in certain motive states. The preparedness of an animal presumably can differ across strains and species. Bolles (1970) demonstrated that experimental manipulations such as CS termination may have quite
different effects for different response modes like running or bar pressing. He maintained that a set of responses, the species-specific defense reactions (SSDR), is emitted in an avoidance situation. If the experimenter-defined correct response is not a member of the set of SSDR's, then the course of learning may be quite tortuous and variable. Since psychologists themselves are becoming aware of the importance of task-specific abilities, it would be pointless for students of behavioral biology to proceed to test the sweeping generalities of dead theories with genetic experiments.

**Sensory Capacities and Preferences**

Among the various processes which are necessary to allow learning to be demonstrated, sensory input obviously occupies a position of primacy. Information must enter the brain before it can be evaluated and stored. Genotypes which lead to differential abilities to gather sensory data should differ in learning rates as a result.

Research with strains homozygous for retinal degeneration (*rd*) has revealed that visual input is necessary for solving certain tasks but not for others. Strains such as C3H and CBA that have rodless retinas did very poorly on black-white discrimination (Wimer and Weller, 1969), pattern discrimination (Bovet-Nitti, 1969), and bar-pressing to turn on a light (Goodrick, 1967), but they could learn a position discrimination quite well (Alpern and Marriott, 1972). Although C3H mice performed very poorly when a light stimulus was employed (Bovet *et al*., 1968), Duncan, Grossen, and Hunt (1971) have shown that good avoidance learning may occur when the light is replaced by a buzzer stimulus (see also Oliverio, 1967). The CBA/J strain was able to learn rapidly to avoid when the task required jumping onto a large platform but encountered great difficulty when the task required running through a small hole (Wahlsten, 1971). However, the CBA/CaJ subline, which has normal vision, was able to learn both tasks as well as other strains with normal vision (Wahlsten, 1972). That this difference between CBA/J and CBA/CaJ was a result of *rd* became clear when F1 mice of a CBA/J by C57BL/6J cross were backcrossed to CBA/J. Retinal degenerate offspring were not different from normals on jump-out avoidance, but they were greatly deficient at one-way avoidance (Wahlsten, 1972). Thus, many of the perplexing results of different experimenters (Bovet *et al*., 1969) may occur only when blind mice are required to run through a small hole in response to a visual stimulus. Although these results should surprise no one today, the presence of *rd* was certainly a source of much confusion in the past, and it impeded progress in the genetic analysis of learning.

Albinism is no stranger to learning research. Lashley (1930) long ago demonstrated that the visual acuity of hooded rats exceeded that of albinos. More recent studies with mice have examined the effects of the *c* gene unconfounded with other genetic differences between strains. When placed
upon a random, segregating genetic background, albinism led to reduced levels of active avoidance learning (Winston and Lindzey, 1964; Winston, Lindzey, and Conner, 1967), water maze learning with either visual or spatial cues (Werboff, Anderson, and Ross, 1967), and straight alley running for food (Tyler, 1970). Albino mice were superior, however, at inhibitory avoidance learning (Winston, Lindzey, and Conner, 1967). Albinism on the isogenic C57BL/6J background was shown to reduce learning of a black-white water maze discrimination (Fuller, 1967) and jump-up avoidance (Henry and Schlesinger, 1967).

Wilcock (1969) recently reviewed these various experiments and concluded that effects of albinism upon behavior are instances of trivial pleiotropy, because lack of eye pigment leads to suppression of nearly any active behavior under bright lights. Several studies have shown that behavioral differences between albino and pigmented mice are greatly reduced when a very dim light is employed over the test area (McReynolds, Weir, and DeFries, 1967; Thiessen, Lindzey, and Owen, 1970). In all of the above studies of albinism and learning which reported illumination conditions, the lights were quite bright, although precise values were never given by the experimenters. Wilcock estimated that they ranged from 50 to 180 ft-c, which is far in excess of levels found to suppress activity in an open-field (McReynolds et al., 1967). Therefore, the albinism effect may have nothing to do with central nervous system differences.

Wilcock's interpretation is supported by a recent experiment by Wahlsten (1972). The albino strain A/J was observed to learn very slowly compared to pigmented strains even under dim red illumination. Mice from an F1 cross of A/J and C57BL/6J were backcrossed to either A/J, C57BL/6J-cJ carrying an albino mutation or albinos from a heterogeneous population; all backcrosses yielded half albino and half pigmented offspring. In no group were albino mice inferior to their pigmented littermates on either jump-out or one-way avoidance learning. Thus, when dim red light is employed, albinism has no effect upon avoidance learning.

Other interpretations of the causes of learning deficits resulting from homozygosity for the albino gene have not been convincing. Fuller (1967) proposed that, since albinism results in a deficiency in both tyrosinase and dopa oxidase, learning deficits might be attributable to an imbalance in brain catecholamines. However, it is known that norepinephrine and related compounds are derived from tyrosine, not via tyrosinase, but rather via the enzyme tyrosine hydroxylase, which functions primarily in nervous system tissue (Cooper, Bloom, and Roth, 1970).

The gene short-ear (se) has been shown to raise the hearing intensity threshold (Bundy, 1951). Denenberg, Ross, and Blumenfield (1963) found no effects of se upon several behaviors, including shock-escape learning. Abeelen (1966) subsequently reported that shock escape learning during jump-up avoidance training was significantly retarded for se/se mice compared to
normal (se/+) littermates; no difference was observed for avoidance learning itself. No reason for the difference was evident.

The above studies indicate that rd and c effects upon learning are indeed trivial when unintended. They leave entirely unexplored the extent of sensory differences between strains, both in terms of relative acuities within a sensory mode and in terms of preferences for one sensory mode over another. Of course, such tests of sensory acuity and preference are time-consuming and require sophisticated learning paradigms. Nonetheless, they could be edifying.

Several reports have appeared of differences in sensory processes between the Tryon rats. Tryon (1940) carried out numerous experiments which showed that surgically disrupting the senses had little effect on the behavior of Brights. Krechevsky (1933) tested Bright, Dull, and unselected rats on his insoluble hypothesis apparatus and observed that the Brights preferred spatial hypotheses, the Dulls used visual hypotheses, and the unselected rats showed no preference. Since these were the only differences noted, Krechevsky attributed the Bright-Dull difference to a "specific response ability" difference. A similar conclusion was reached by Wherry (1941), who subjected various response measures on the Tryon maze to factor analysis; the scores of Brights and Dulls on his three factors, forward going, food pointing, and goal gradient, suggested that Brights showed spatial and Dulls visual orientations. Later work indicated that S_1 (Bright) were superior to S_3 with spatial cues but not with visual cues (Markowitz and Sorrells, 1964; Markowitz and Becker, 1965). However, Fehmi and McGaugh (1961) reported that S_1 was superior on a more difficult horizontal-vertical discrimination, which certainly required the utilization of visual cues. Although sensory abilities and preferences are indicated, conclusive evidence of their relevance to maze learning differences between the two lines is lacking.

**Motivation**

The relation between motivation and learning has a long history of theoretical dispute (see discussion by Kimble, 1961, Chap. 13). One central issue concerns the necessity of proper motivation to assure learning at all. Unfortunately, demonstrations of latent learning, sensory preconditioning, and transfer between drive states have not been attempted with genetic experiments.

Whereas diverse opinions exist concerning the need for motivation to assure the acquisition of information, most theorists recognize the importance of proper motivation in order to guarantee the reliable performance of a learned response (see Estes, 1969). Vast research indicates that simple, unitary responses are acquired more rapidly when the animal is more highly motivated by either food or water deprivation or electric shock (Bitterman and School, 1970). Hence it would surprise no one if strains found to learn at different rates also were differentially motivated by identical operations or if motivation changed as a correlated response to selection for learning rate. Of course,
neither would it be surprising if motivational differences accounted for only part of the variation in learning rates. Pure associative learning ability might vary as well. The problems are complicated by the observation that more complex tasks appear to have an intermediate level of motivation for optimum learning; a simple increase in motivation may actually lead to poorer learning of complex mazes or shuttle avoidance (Bitterman and Schoel, 1970). The only way to determine these contributions is to measure motivation independently from the learning task of interest. If the operations which yield equivalent states of motivation in various strains can be determined and then applied in training, differences in learning rate beyond motivational differences may be determined.

Using inbred mouse strains, Carran, Yeudall, and Royce (1964) demonstrated that large differences in shuttle avoidance for C3H, C58, and SWR mice at low-shock voltage disappeared entirely at higher voltages. Likewise, C3H showed greater passive avoidance than C58 mice at all but the greatest pressure of air blast (Carran, 1967). Although their results suggested that motivational variation existed, they did not establish that learning differences at lower shock or air blast levels were caused by motivational differences. Wahlsten (1971) addressed this problem by training with shock levels which equated the unconditioned response to shock for several strains. The amount of jumping and squealing to six intensities of shock was determined for four inbred strains. Then the shock intensity was calculated which gave for each strain the same amount of jumping as for the average of the strains at 180 μA (5.63 jumps/4 sec of shock). Training naïve mice on a jump-out task with shocks which equated jumping in the pretest totally eliminated between-strain variation in latency of the first escape but did not substantially modify the magnitude of variation in learning rate as compared with training at 180 μA. Training with one-way or optional (either jumping or running allowed) avoidance also suggested little or no relation between initial response to shock and rate of learning. Data on two F1 hybrids and a four-way cross suggested that the mode of inheritance of the two measures was different; only learning rate exhibited significant heterosis. Since the frequency of jumping may not be a perfect indicant of motivation during shock, motivational differences cannot be ruled out entirely.

Selection for learning has produced motivational differences in two instances. As mentioned above for the Tryon strains, the Brights appeared to be more highly motivated by hunger, while the Dulls had greater aversion to water (Searle, 1949). Variable results obtained with shock motivation. Heron's (1935) rats were selected on a maze task very similar to Tryon's. When Heron and Skinner (1940) extinguished bar-pressing for food reward, they found that more rapid extinction for the maze dull strain could be attributed to its lower rate of pressing at the onset of extinction; they suggested that the brights were more hungry. Harris (1940) reanalyzed the original Heron maze data and discovered that the ratio of running time to mean errors on a trial was
generally smaller for the brights, which was held to be indicative of a weaker drive state in the dulls. Kruse (1941) observed that the brights ate more food under the usual deprivation condition and that they seemed to be more emotional, too. In these respects, Heron's rats resembled those of Tryon, for mild motivational differences were noted in both groups. In neither case were the motivational differences proved to be genetically related to learning differences.

The McGill bright and dull rats selected on the Hebb-Williams maze (Thompson, 1954) have also received some attention. It is interesting to note that a prime reason for using the Hebb-Williams battery of problems was to select for a more general learning ability and thereby circumvent the "less interesting" motivational differences produced by Tryon and Heron. When Thompson and Bindra (1952) tested the F4 generation of selected rats for food eating, eating time, defecation, urination, and timidity, a significant strain difference was obtained only for urination. Thompson (1953) also tested exploratory activity under several deprivation levels, but again no strain differences were manifest. Thus, Thompson's original goal was met; learning differences existed without concomitant motivation or emotion differences. Unfortunately, the McGill strains have not been the subjects of extensive learning tests as were Tryon's.

Memory

The ability to retain as well as store information is obviously a prerequisite for successful retrieval of that information at some later time. An animal of a certain genotype which either fails to store information permanently or stores it in a manner that makes retrieval difficult would appear to be deficient in acquisition of any task. Evidence exists that the process of memory storage requires a certain amount of time before a permanent record is made (McGaugh, 1966); the memory becomes less susceptible to disruption by diverse insults as time progresses. Thus, a strain which has a slower rate of memory "consolidation" would appear to be retarded in acquisition of a task at a fixed intertrial interval, assuming the interval is considerably shorter than the time required for efficient storage. Likewise, a strain which could not enter information into long-term storage at all would appear to be grossly deficient with widely spaced trials.

The work by McGaugh and his colleagues has shown that the Tryon S1 and S3 strains differ in the time-dependent aspects of memory storage but that both strains are able to enter information properly into long-term storage. The spacing of trials on a Lashley III maze was important, for the superiority of S1 at short intervals (ITI 30 sec) vanished at an ITI of 5 min or more (McGaugh, Jennings, and Thomson, 1962). A later study (McGaugh and Cole, 1965) found that ITI interacted with age, for in young rats S1 was superior only at a long ITI (30 min). The difference between S1 and S3 with massed
trials was eliminated by pretrial injection of the drug 1757 I.S., which improved learning only for S3 (McGaugh, Westbrook, and Burt, 1961). Spaced trials (one per day) gave equivalent performance for S1 and S3 on a 14-unit T maze, and posttrial injection of picrotoxin greatly facilitated learning by S3 only (Breen and McGaugh, 1961). These studies supported the hypothesis that the rate of consolidation was normally faster for S1 but could be accelerated in S3 by administering stimulant drugs. This notion was strengthened by a study of the time-dependent effects of posttrial ECS using a Lashley III maze and one trial per day (Thomson et al., 1961); if no ECS was given, errors by S1 and S3 were equal, but ECS 45 sec after a trial increased error scores more for S3 than S1 and ECS at 75 sec increased errors above control levels only for S3. Similar facilitation of learning by posttrial injection of physostigmine for S3 but not for S1 on a Lashley III maze was reported by Stratton and Petrinovich (1963), but they observed a large difference in favor of S1 in the control group at one trial per day, which contradicted the finding of Thomson et al. (1961). Perhaps this can be attributed to their learning measure, trials to criterion, which differed from the usual procedure of giving a fixed amount of training. Although the experiments did not prove that the original Tryon strains diverged in learning rate because selection produced memory differences, McGaugh's research leaves little doubt that the S1 and S3 strains differed in memory processes. The differences were of such a magnitude as to account for virtually all of the between-strain variation in acquisition rate. Perhaps the most important implication of this finding is the extent to which memory processes are determinants of learning ability. In fact, only recently have learning theorists given due consideration to memory processes (see Estes, 1970).

Other research on genetic differences in memory is less convincing. Bovet, Bovet-Nitti, and Oliverio (1969) presented data which showed that retention of a single passive-avoidance experience was good 10 sec after training but poor 24 hr later for C3H/HeJ mice; the reverse was obtained for DBA/2J mice. In addition, short intertrial intervals in shuttle avoidance led to good learning within a session for C3H mice, but retention was poor 24 hr later. On the other hand, DBA mice showed less change within a session but excellent retention the next day. The various experimental results led Bovet et al. (1969) to suggest that C3H and CBA mice have good short-term memory (STM) but poor long-term memory (LTM) while DBA mice have poor STM and good LTM. A most unfortunate aspect of their work was their concentration on two inbred strains, C3H/HeJ and DBA/2J, which differ in numerous ways other than learning ability. Recent evidence has demonstrated that strain differences are quite small when tasks are employed that do not require the utilization of visual cues by C3H mice with retinal degeneration. The strains C3H/HeJ, CBA/J, and DBA/2J all show good short-term retention of a simple active avoidance task (Wahlsten, 1971, 1972). Both C3H and DBA also appear to have intact long-term retention for several avoidance tasks (Duncan,
Grossen, and Hunt, 1971; Wahlsten and Weening, unpublished data). Results of other researchers reporting memory differences in mice (Wimer et al., 1968; Randt et al., 1971) must also be viewed with skepticism because they showed that DBA/2J exhibits poor long-term retention, which is contrary to the data of many others.

Other studies which involved tests of long-term retention in many strains found no significant strain differences (Henderson, 1968a; Stasik, 1970). Thus, certain inbred strains of mice may have impaired long-term retention or retarded consolidation rates, but their identities are currently unknown.

If memory variations underly differences in learning rate for a wide range of strains besides S1 and S3, it will be necessary for research of a magnitude similar to that of McGaugh’s to be undertaken. The importance of memory processes will be underscored if, for example, Bignami’s RHA and RLA strains show time-dependent differences as well.

Emotionality

Animals which are otherwise quite capable of efficient learning may perform very poorly if a particular training situation evokes strong competing responses. In an avoidance learning task, freezing may appear to be a concomitant of great “fear” or “emotionality,” or it may be learned because of unforeseen reinforcement contingencies which encourage freezing (McAllister and McAllister, 1971). Wilcock and Broadhurst (1967) obtained measures of defecation and ambulation in an open field, a presumed test of emotionality, in five inbred rat strains and then trained them in shuttle avoidance. The Pearson correlation between mean open-field defecations and mean number of avoidances for each strain was +.06, which hardly supported any interpretation of the emotionality hypothesis. Reynierse (1970) has performed several experiments which suggest that rats of the Sasco strain are more emotional and extinguish avoidance responding more quickly than Holtzman rats under certain conditions. However, in no experiment was a strain difference in rate of initial acquisition observed under any duration of safe compartment confinement. A conflict situation was shown to decrease the learning of shuttle avoidance by BALB/c mice but to have no effect upon relearning by C57BL/10 mice (King and Mavromatis, 1956); an increase in freezing, a presumed concomitant of high emotion, was reported for the BALB strain. Skin-resistance changes resulting from electric shocks, which were believed to indicate relative fearfulness (Carran et al., 1964), were used to explain why the more “fearful” (i.e., greater resistance decrease after shock) C3H mice were better at both active shuttle (Carran et al., 1964) and passive avoidance (Carran, 1967). Fuller (1966) trained three strains of mice on Sidman avoidance in a shuttle box after injection of several doses of the tranquilizer chlorpromazine. While the rate of responding decreased for all strains at higher doses, the effect was minimal for the RF strain but quite large for C3HeB and
C57BL/6 animals. Since the RF strain had a much lower operant rate than the other two under the placebo condition and showed little drug effect, it may have had a lower level of fear or emotion. Thus, in two experiments highly emotional animals learned to avoid more quickly or proficiently, in one experiment the highly emotional strain performed more poorly, and in two others there was no relation between "emotionality" and avoidance acquisition.

Similarly, the failure to modify open-field defecation by selection for shuttle avoidance learning (Broadhurst and Bignami, 1965) contrasts with the significant differences in shuttle avoidance obtained after selection for open-field defecation (Broadhurst and Levine, 1963).

These difficulties may be attributable in part to previous measures of emotionality. Low activity, as indicated by few square crossings in an open field, has generally been held to indicate freezing or immobility, but direct observation of postures of several inbred and selected mouse strains has revealed freezing to be a very rare event (Streng, 1971); mice with low activity scores tend to spend more time "air sniffing" or "object sniffing." The open-field defecation measure seems to be related more to social dominance or territorial marking than to fear in some situations (Bruell, 1969; Brain and Nowell, 1969). Other evidence suggests that rate of responding in avoidance training may be more clearly related to "kinetic drive" than to fear or emotionality (Fuller, 1970). Thus, further research on genetic variation in emotionality and avoidance learning must await the development of more meaningful operational definitions of emotion or fear. One promising approach would be to measure directly the competing responses by observation or photographic analysis.

The Nervous System

Since learning is presumably a manifestation of the functioning of the brain, strains whose brains differ radically should likewise differ in learning ability. The big question here, though, is which of the multitudinous aspects of the brain are related to learning.

The weight of the brain appears to bear little or no relation to learning ability in rats and mice. Brain weight-learning correlations have been inconsistent over the years for the Tryon S1 and S3 strains (Rosenzweig, 1964). Furthermore, the brains of Heron's bright and dull strains did not differ in weight after 14 generations of selection (Silverman, Shapiro, and Heron, 1940). Wimer and Prater (1966) found that mice selected for high brain weight required fewer trials to learn a black-white discrimination than those selected for low brain weight. However, Collins (1970b) found that the largest difference in discrimination learning was not between high and low lines but instead was between the control line (more errors) and the selected lines. In addition, environmental enrichment or isolation had different effects on brain
weight and learning ability of the selected lines (Collins, 1970b). Although brain weight-learning correlations have not been reported for inbred mouse strains, comparison of strain variation in brain weight (Wimer, Wimer, and Roderick, 1969; Wahlsten, Hudspeth, and Weening, unpublished data) to strain differences in avoidance learning of several studies (Table 1) reveals no consistent rank correlation.

The chemistry of the rat brain has received much attention with regard to learning. Studies of the Tryon S1 and S3 strains and Roderick's high and low AChE strains have revealed that the absolute concentration of single neurotransmitters does not correlate highly with maze learning (Rosenzweig, 1964). However, the relative concentration of ACh and AChE suggests that rats with higher ACh/AChE ratios are better able to solve mazes. As Rosenzweig himself pointed out, the data are not conclusive, and more research with other strains is needed. Nonetheless, the important idea that study of the joint functioning of many important neurochemicals is required to understand learning should be manifestly clear.

Abundant research on genetic variation in the chemistry of mouse brain has been reported (Sudak and Maas, 1964; Schlesinger and Griek, 1970), but observed differences have not been related to learning ability.

Structural and organizational attributes of the brain have received scant attention in the genetic context. Wimer et al. (1969) found that inbred mouse strains which had a neocortex of relatively large volume tended to have a hippocampus of relatively small volume (Spearman r = -.83). They did not attempt to relate their data to learning ability. Visual pathways have been found to differ dramatically in albino and hooded rats (Lund, 1965). The organizational differences related to patterns of interocular transfer (Sheridan, 1965) and visual evoked potentials (Creel, Dustman, and Beck, 1970). Their relevance to normal learning differences has not been established. Neither have they proved that the differences are caused by the gene c in random bred populations.

Given the large number of mutant genes which are known to affect brain organization (Sidman, Green, and Appel, 1965), it is likely that alleles more within the normal range of variation have similar effects upon organization. Future studies which examine detailed organizational aspects of brain, instead of homogenizing these differences, may detect patterns which relate to learning ability.

Discussion of Genetic Correlates

The above studies of genetic correlates of learning emphasize several points mentioned earlier.

1. Presentation of mere correlations between learning ability and other attributes of inbred strains cannot establish a causal relationship. The strains must be crossbred, and the correlations must be observed in segregating generations.
2. Since imperfect relationships are to be expected, the actual magnitude of the genetic correlation between phenotypes should be computed. Environmental sources of phenotypic correlation can be similarly derived.

3. A causal relationship should also be demonstrated by independently manipulating the variable of interest and then observing consequent changes in learning rate.

Future progress in the study of genetic correlates may also be expected if single genes which modify learning ability are isolated. As mentioned above, the pathways of two genes known to modify learning rates end at the periphery, the eyes for albinism and retinal degeneration. Future studies may uncover highly informative pathways, but the risk of further trivial outcomes is high because of the grossly deleterious effects of most known mutant genes. Perhaps attempts to identify and map single genes which specifically affect learning, as Collins (1970a) has done for audiogenic seizures (audiogenic-seizure-prone gene asp), will provide exciting results. In fact it would be worthwhile to examine pleiotropic effects of asp on learning.

It would also be worthwhile to conduct a careful, large-scale program of artificial selection for learning, since all of the past selection experiments had one or more serious flaws. The availability of such excellent genetic material might enhance the chances of some researchers discovering important correlates of learning. The best selection criterion to use for such an experiment is not clear, however. Some researchers would certainly favor general learning ability by selecting for a pooling of an individual's scores across several diverse learning tasks which encompass a wide range of stimuli, responses and motives. Others might prefer lines selected for rate of learning a simple task such as a T-maze. The latter procedure would allow the analysis of all of the components of the learning process, from sensation to motivation. Of course, conducting both experiments would yield the most information.

ADAPTIVE VALUE OF LEARNING

The fundamental theorem of natural selection asserts that the rate of increase in fitness in a population is equal to the amount of additive genetic variance of fitness at that time; after many generations of selection, those characters most closely related to genetic fitness will reach their maximum mean level in the population and will have no remaining additive genetic variation. Roberts (1967b) has suggested that phenotypes which exhibit high heritability may not be very important determinants of "fitness" in the genetic sense, because traits which determine biological fitness tend to have very little additive genetic variation. For example, in cattle the amount of white spotting in the coat has a heritability of .95, while conception rate has a heritability of only .01 (Falconer, 1960, p. 167).

Examination of Table 3 reveals that heritabilities of learning phenotypes range from low (.2) to moderate (.5). How these values relate to fitness in the
genetic sense cannot be known unless heritability of known fitness characters, such as fertility or litter size, are calculated for the same populations. Oliverio et al. (1971) directly compared $h^2$ of learning and wheel-running phenotypes, but no researchers studying learning have shown interest in reproductive abilities of their subjects.

It is possible that certain categories of learning may have different adaptive values and therefore different patterns of inheritance than others. Low heritability for determinants of fitness may have an interesting relation to the hypothesized low heritability of simple learning tasks. It is quite conceivable that life in the wild imposes a high premium on learning quickly which things are nutritious and which are nasty but does not discriminate among levels of ability to solve intricate multiple-contingency tasks with high information content. This outcome should occur if higher mental abilities are not necessary to solve most of the problems of survival. The adaptive value of learning, and hence its heritability, may also be related to the breadth of a species' ecological niche. High ability to store and retrieve information should be especially useful when an animal typically encounters a wide range of foodstuffs, competitors, and building materials. Animals which occupy a very narrow niche, on the other hand, may be able to solve most problems with stereotyped responses to a limited number of stimuli. The ecological niche may also influence the kinds of learning abilities which will be highly developed in a certain species (see excellent discussion by Eibl-Eibesfeldt, 1970, Chap. 13).

Another attribute of traits with high adaptive value is that they tend to degenerate during inbreeding and show a great increase when inbred strains are crossed. Since natural selection acts to reduce additive genetic variation by eliminating the less fit genotypes, the only genetic variance remaining after many generations of selection for traits closely related to fitness should be attributable to heterozygote superiority. This means that components of fitness should exhibit overdominance as well as low heritability. This important principle allows one to distinguish between traits having low heritabilities because of sloppy measurement or other causes of a large $V_E$ and traits which are major components of genetic fitness.

Simple crosses between strains may be used to detect the presence of dominance effects on learning. When two strains are crossed to form an $F_1$ hybrid population, the average degree of dominance may be determined by comparing the $F_1$ mean score to the mean of the two parent strains, the midparent score (MP), or to the highest scoring parent (HP). All instances where $F_1$ is greater than MP are characterized by hybrid vigor or heterosis. The results of several such genetic studies of learning are summarized in Table 4. In most studies employing inbred mouse strains as parents, significant directional dominance was observed. The $F_1$ hybrids were generally superior to the average of their parents for learning of a two-choice maze for food reward (Vicari, 1929), lever pressing for food reward (Smart, 1970), water-
escape learning (Winston, 1964; Winston and Lindzey, 1964), shock-avoidance learning (Collins, 1964; Schlesinger and Wimer, 1967; Abeelen, 1966; Rose and Parsons, 1970; Wahlsten, 1971; Oliverio et al., 1971), and CER conditioning (Henderson, 1968a). Many instances of overdominance were also reported.

Several experiments with selected strains, summarized in Table 4, have been reported. Neither Tryon (1940) nor McGaugh, Westbrook, and Burt (1961) found heterosis in a cross of the Tryon bright (S₁) and dull (S₃)

### Table 4

Comparisons of F₁ Hybrid Scores with Mid-Parent (MP) and High-Parent (HP) Scores

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of strains</th>
<th>Parent</th>
<th>F₁</th>
<th>F₂</th>
<th>F₁&lt;MP</th>
<th>F₁&gt;MP</th>
<th>F₁&gt;HP</th>
<th>Heterosis?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inbred parents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vicari (1929)</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>Collins (1964)</td>
<td>5</td>
<td>20</td>
<td>0</td>
<td></td>
<td>5</td>
<td>1</td>
<td>14</td>
<td>Yes</td>
</tr>
<tr>
<td>Winston (1964)</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td></td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>Winston (1964)</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>Winston and Lindzey (1964)</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td></td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>Abeelen (1966)</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Schlesinger and</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Wimer (1967)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henderson (1968a)</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td></td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>Rose and Parsons (1970)</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td></td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>Smart (1970)</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Stasik (1970)</td>
<td>6</td>
<td>15</td>
<td>0</td>
<td></td>
<td>12</td>
<td>1</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Oliverio et al. (1971)</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>Oliverio et al. (1971)</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Wahlsten (1971)</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td></td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Noninbred parents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tryon (1929)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>McGaugh et al. (1961)</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>Bignami (1965)</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>Fuller and Scott (1954)</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>Scott and Fuller (1965)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>No</td>
</tr>
</tbody>
</table>

*a* Control condition only.

*b* Infantile trauma condition.

*c* Suppression ratio over eight trials on the second day of CER training.

*d* "Efficiency" of performance on several schedules of partial reward.

*e* Shuttle shock-avoidance learning.

*f* Flashley III maze learning for food reward.

*g* Jump-out and one-way avoidance task for each of the F₁'s and F₂ (four-way cross).

*h* Control condition only.

*i* Three tests on same F₁ dogs.

*j* Same three tests as (i) above, plus two additional tests.
strains. In both studies the F₁ mean was very close to MP. Bignami (1965) obtained moderate heterosis in a cross of his high (RHA) and low (RLA) avoidance strains taken from the third generation of selective breeding. The mean numbers of avoidances in 250 trials were 170.9 for RHA, 46.1 for RLA, and 143.7 for their two reciprocal crosses, which was greater than MP (108.5) but less than HP (170.9). Bignami's data suggest that only a moderate degree of directional dominance existed.

There are also a few reports of crosses between strains of dogs, which were known to be similar but still possess genetic variation. Consistent directional dominance was not observed in any study.

The lack of detectable heterosis with heterogeneous or selected strains does not contradict the positive results obtained from inbred strains, for the F₁ mean will result from additive as well as dominance causes when the parent strains have genetic variation. Only when isogenic parent strains are employed will the F₁ versus MP difference reflect dominance effects alone (Bruell, 1967). In fact, the above studies confirm the notion that hybrid vigor is the precise opposite of inbreeding depression, because heterosis is obtained only if extreme inbreeding has occurred previously. A well-known effect of inbreeding is to eliminate heterozygosity (Falconer, 1960). Thus, these studies also point to the importance of dominance as a genetic mechanism which influences learning.

One difficulty with this simple dominance explanation of hybrid vigor arises when parent and F₁ variances are compared. Since F₁ of a cross between two highly inbred strains has no genetic variance, the phenotypic variance should not differ significantly from that of the parent strains. If the variances differ, significant epistatic interaction between loci probably is involved (Mather, 1949). Although Winston (1964) found that F₁ variances resembled those of their parents, Schlesinger and Wimer (1967) observed a substantial reduction in the variance of most F₁ hybrids. The most extreme case was a cross of DBA/2J and C3H/HeJ; the standard deviations in trials to acquisition were 8.37 and 9.33 for DBA and C3H, respectively, and 1.4 for F₁. The reduction in F₁ variance was of a magnitude similar to several examples given by Falconer (1960, Table 15.2). Rose and Parsons (1970) noted reduced variability in a learning score for F₁ hybrids only early in training.

Another problem appears in studies of dominance variance in heterogeneous populations. Significant dominance variance will lead to an intraclass correlation between full-sibs which is more than twice that between half-sibs in sib analysis (see Falconer, 1960). However, applications of sib analysis to learning (Table 3) have found no evidence of dominance variance (Willham et al., 1963; Oliverio, 1971; Oliverio et al., 1971). This was somewhat unexpected in the experiment of Oliverio et al. (1971), since substantial dominance was indicated in the crosses of inbred strains from which the randomly bred populations were derived. These results also suggest that epistasis may be important.
Thus, neither of the criteria for inheritance of fitness characters, low heritability and heterozygote superiority, are unequivocally met by current data on the learning phenotype.

Another problem for the study of the adaptive value of learning is that genetic research has been conducted in the lab with domesticated animals. Lab strains have undergone selection as well as inbreeding since being rudely snatched from their feral homes. Whether their genetic composition resembles that of their ancestors (whose offspring presumably are still afield) thus becomes an empirical question (see Bruell, 1967).

The means by which these difficulties may be overcome are quite numerous. Study of learning ability of wild populations would be a good place to start. Although methodological problems are certain to be encountered in the study of truly wild animals, transporting them to seminatural habitats which allow controlled observation and stimulus presentation as well as individual identification might provide a good starting point. Commensal populations, which already live in close proximity to man, are especially good candidates for such experimentation (Bruell, 1970; Selander and Yang, 1970). It would be important to test the animals before too many generations had elapsed away from the original environment.

Another strategy of immediate utility would be to release groups of lab animals of known gene frequencies and learning abilities into environments in which only the influx of migrants of the same species was controlled. Subsequent generations could be retrieved, "domesticated," and then tested for learning and so forth. Environments could be arranged with and without predators or with and without a limited food supply. This strategy would be especially interesting if strains of animals selected for either high or low ability to learn certain kinds of tasks were to be released into seminatural environments and their abilities to adapt to various conditions were then to be observed.

Although such efforts require substantial time and effort, they must be undertaken in order to discover the true function of learning ability for the individual and for the population.

GENOTYPE-ENVIRONMENT INTERACTION

The phenotypic expression of a particular genotype is known to reflect the individual's postfertilization environment prior to the time of testing. The important question in this regard is whether genotypes which lead to superior learning in one environment will be similarly endowed across a wide range of living conditions. If genotypic and experiential components of learning ability are truly additive \( P = G + E \), then conclusions drawn from studies of limited scope may be expected to have broad validity.

The experiment by Cooper and Zubek (1958) demonstrated that rearing Thompson's (1954) bright and dull rat strains in either an enriched or an
impoverished environment eliminated the strain differences in learning that were originally produced by selection in a normal lab environment. Likewise, pretraining experiences have been shown to affect some standard strains more than others. The handling of infant rats did not change later shuttle avoidance learning of the Sprague-Dawley strain, whereas handling greatly improved subsequent avoidance of both the Harlan and Rockland Long-Evans strains (Levine and Wetzel, 1963); with infantile handling Sprague-Dawley and Rockland were equivalent, while Sprague-Dawley was superior under the unhandled control condition. Infantile trauma (loud noises) increased the number of errors on later learning of a four-unit T maze equally for the three strains of mice tested by Winston (1963). Lindzey and Winston (1962) reported that gentle stroking before a trial improved learning of a six-unit T maze for the C57/B1/1 strain but did not change the scores for C3H/Bi. Freedman (1958) reported that either indulging or disciplining puppies of four strains of dogs had very temporary differential effects upon later inhibition training. Thus, early experience has highly variable effects on the learning abilities of different genotypes.

Experiences prior to training may also affect the expression of hybrid vigor. Winston (1964) observed that infantile trauma, a loud noise, increased the number of errors in a water-escape maze for inbred mice but had minimal effects upon the F₁ hybrids. One consequence of this operation was that all hybrids were superior to HP in the trauma condition, whereas only one of three hybrids exhibited any heterosis at all under the control condition. Henderson (1970) has recently shown that a restricted early environment can greatly reduce the differences between inbred and hybrid mice on a complex exploratory task. Hence, not only may hybrids be less affected by trauma, but they may also benefit more from varied experience in an enriched environment.

The potential complexity of genotype-environment interaction increases as more strains are raised in more different environments and are then tested on several learning tasks. Henderson (1968b) reported preliminary results of a diallel cross of six inbred strains reared in either a standard or an enriched environment and then tested on six different learning tasks. The results indicated that "...there was little consistency in which genotypes benefitted most from enrichment with respect to each of the learning tasks..." (p. 149).

It is apparent that learning phenotypes are subject to a multitude of complex genotype-environment interactions. While these results certainly tend to obfuscate and frustrate our attempts to discover general principles of the inheritance of learning ability, they also are important facts about the learning process. If future research is able to discover the basis for these interactions, our understanding of learning will increase many fold.
CONCLUSIONS

Of the various questions discussed above, only one, the null hypothesis, has been answered.

The question of the relative magnitude of genetic variation can be viewed as somewhat ill-conceived. Since heritability can vary as a function of so many conditions, it is hoped that any visions of a true, invariant estimate have vanished. Further studies to measure heritability of a particular learning phenotype in laboratory populations would appear pointless.

The degree to which learning ability has adaptive value cannot be determined until populations are studied in which the multifarious forces of natural selection are allowed to apply unfettered. Although psychologists implicitly assume that learning ability has great utility for animals, the maintenance of high heritability of learning under natural conditions would imply that learning has really little relation to fitness.

Certainly the most important problem in future research will be to identify the genetic correlates of learning. Virtually nothing is presently known about the physiological bases of genetic differences in learning. The pathways of major genes affecting learning ability in the normal range are likewise unexplored. This situation is surprising in view of the great efforts that neurobiologists make to modify the learning rates of animals or to compare widely divergent species whose differences can never be subjected to genetic analysis. Animals of different learning abilities are readily available that have never endured electrical devastation or psychopharmacological perdition.

Genetic methods may also be applied to some of the major questions within the areas of learning and memory research. Controversy over the unitary or dual nature of certain processes is particularly susceptible to genetic clarification. For example, it is of interest to know whether classical and instrumental learning are two distinct processes or different reflections of the same basic learning process (Miller, 1969; Rescorla and Solomon, 1967). If a situation can be devised in which classical and avoidance training are administered with identical CS, US, and response mode to different members of parent and offspring generations, it would be possible to calculate the genetic correlation between learning under the two contingencies. A very high $r_A$ would indicate that they in fact depend upon the same process, while $r_A = 0$ would suggest that they are essentially independent processes. Intermediate values of $r_A$ would mean that the processes share common elements but also have unique aspects. Similar experiments can be done to study the similarities of short and long-term memory as well as motivation and "pure associative" learning ability. Quantitative genetic analysis is especially useful in answering these questions because it is entirely empirical (does not require an
hypothesis) and can detect a wide range of possible outcomes with predictable accuracy.

The study of genetic differences in learning will most likely lead to some important discoveries about the mechanisms of learning, but it can never be relied upon to identify all of the important variables. All of the genes which contribute to learning differences can be identified, at least in principle, but all the genetic loci which are fixed for one allele in a certain population will remain undetected, even though they may mediate crucial processes in the storage and retrieval of information. This is true because genotypes are inferred from knowledge of phenotypes. If only one allele occurs at a particular locus, there will be only one genotype, and hence all animals will be affected similarly. In fact, the process of natural selection will tend to produce genetic uniformity at those very loci which are most important for adaptive behavior. Whatever genetic variation does exist may be "permissible" variation which, nonetheless, leaves the most important components of the learning process inviolate. Suffice it to say that within the foreseeable future this limitation will probably be the least of our difficulties.

REFERENCES


Henderson, N. D. (1970). Genetic influences on the behavior of mice can be obscured by laboratory rearing. J. Comp. Physiol. Psychol. 72, 505-511.


Wahlsten, D. (1972). Contributions of the genes albinism (c) and retinal degeneration (rd) to a strain-by-training procedures interaction in avoidance learning. Unpublished manuscript.


**SUPPLEMENTARY REFERENCES**


