Quantitative Genetics Project
Report
1947 - 1957

R. E. Comstock, H. F. Robinson and C. C. Cockerham

Institute of Statistics
Mimeo Series No. 157
# Table of Content

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II. Summary and Discussion</td>
<td>4</td>
</tr>
<tr>
<td>III. Theory</td>
<td>28</td>
</tr>
<tr>
<td>IV. Investigations with Corn</td>
<td>54</td>
</tr>
<tr>
<td>V. Experimental Results with Crops Other Than Corn</td>
<td>92</td>
</tr>
<tr>
<td>VI. Listing of Publications</td>
<td>97</td>
</tr>
</tbody>
</table>
I. Introduction

A. Perspective

The general objective of this research program has been to obtain useful information about the genetics of quantitative characters in crop plants. Throughout, the problems of plant breeding have provided our points of departure. This implies no lack of interest in fundamental aspects; the best and ultimate solutions of breeding problems require basic information. However, our interests in plant improvement have affected the pattern of our research. We have deliberately chosen to investigate issues having obvious immediate implications in plant breeding. At the very outset our interests tended to be dominated by those immediate implications but we think it fair to say that as our own insights deepened we have managed to place appropriate emphasis on critical basic issues without compromising our interests in the immediacies of plant breeding. That this should be possible is obvious. Practical issues hinge inevitably on basic facts. Nevertheless an awareness of the semi-dual nature of our objectives will enable a fuller appreciation of the following report by helping to explain directions taken in our research and our tendency to view results in two frames of references.

B. General outline of the research

The total research program is outlined in terms of a theoretical and an experimental phase. The latter is further subdivided into work with (a) corn and (b) other crops. The work with corn has been conducted by members of the Genetics group of the Experimental Statistics Department. With only two exceptions all work with other crops has been conducted by other plant geneticists of the North Carolina Agriculture Experiment Station and our participation has been in an advisory capacity.
The experimental and theoretical investigations are intended to be complementary. The experimental program is designed to provide new evidence on timely issues whereas objectives in the theoretical phase are to identify key issues, to devise and evaluate experimental approaches, to extend the bases for interpretation of data, and to correlate in an interpretive way the results of the whole experimental program in the light of past and current literature.

Intimate connections between the various parts of the whole program were inevitable. Theoretical considerations have influenced the course of experimentation, experimental results in turn have raised issues demanding theoretical examination, and cross-meanings are showing up in the evidence from different crops. In the light of these interrelations we believe it desirable to attempt putting all results in a common perspective. This has influenced our organization of this report.

C. Organization of the report

The following objectives have guided our formulation of this report.

(1) To report results in each portion of the program in reasonable and appropriate detail.

(2) To provide a useful synthesis of the evidence from the whole program.

(3) To provide an interpretive summary that is self-contained and readable with minimum reference to the remainder of the report.

Part II of the report represents our effort toward a self-contained synthesis and summary of present evidence. Parts III, IV and V are devoted to results obtained in the separate portions of the program. Discussion in these parts is minimized and the intent was to limit it to what was particularly pertinent where included or had no application at all elsewhere. Appropriate detail in presentation of results was a matter of judgement. Much more could have been given. Admittedly,
our judgement was strongly affected by our preoccupation with what seem to us to be the most significant aspects of results to date rather than sheer volume of data.

Because our primary experimental crop has been corn and our interests in many respects have radiated out from our interests in that crop, we have chosen to center the report around the evidence pertaining to corn. This has affected our reporting of results with other crops in two ways: (1) We have included only those results that in our opinion contribute to interpretation of corn data either because they add to the evidence subject to interpretation on common grounds or because they provide potential contrasts by reasons of specie differences in breeding systems (mode of reproduction). (2) Results included are described in less detail than in the case of corn.

Finally, partly for the sake of brevity and partly a reflection of our dominant interest to date, we have reported results for traits other than yield in only a fragmentary way. In fact what has been presented relative to other traits is offered primarily as a phase of the evidence pertaining to yield; to show significant and meaningful contrasts. If any defense were needed it would revolve around these facts: (1) yield is the dominant economic trait and (2) by reason of being the dominant economic trait and because of its intimate relation to "fitness", grain yield in corn has inevitably been deeply involved in the evolutionary development of the specie with the result that it is probably more significant relative to the over-all genetic status of modern corn than any other character on which data is as easily obtained.
II. Summary and Discussion

This summarizing presentation could proceed along any of various pathways. We have chosen to begin with a compact summary of our own experimental evidence subdivided under the following headings:

(i) Additive genetic variance
(ii) Relative magnitudes of dominance and additive genetic variance
(iii) Variance from interaction of genotype with environment
(iv) Epistasis
(v) Differences between varieties

The discussion to follow will deal more or less simultaneously with the first three of these and separately with each of the last two, in so far as genetic interpretations are concerned. The remainder of the discussion will be devoted to

(i) Implications for corn breeding and
(ii) Implications for future research.

As already indicated, attention will be centered on experimental evidence pertaining to grain yield. A major portion of the evidence relates to the amounts of additive genetic and dominance variance and through these to level of dominance in genes affecting yield.

Our approaches (experimental and theoretical) have always revolved around the categories of gene action (strict additivity, dominance and interaction of non-allele genes) and the related subdivision of the total genetic variance proposed by Fisher (Trans. Roy. Soc. Edinb. 52:399-433, 1918). It is vitally important to note that the Fisherian definitions of additive gene effects and additive genetic variance apply whether or not any or all genes are involved in either (or both) intra-allelic or inter-allelic interactions. That is to say regardless of the interrelations among responses evoked by the different genes; each has, by the definition adopted, an
additive effect though depending on circumstances this effect may take the value zero. The other, and equally important, side of the issue is that unless a gene has no dominance relation with any allele and no epistatic (non-allelic interaction) relation with any non-allelic gene, the additive effect of the gene is a function of one or more of the gene frequencies of the population. The additive effect of a gene is a constant only when the gene is not involved in either intra-allelic or inter-allelic interaction with any other gene. The additive genetic variance is always variable with gene frequency. If dominance is absent at all loci, there is no dominance variance; otherwise the dominance variance is variable with gene frequency. If there are no interactions among non-allelic genes, there is no epistatic variance; otherwise epistatic variance is variable with gene frequency. This variation in dominance and epistatic variance occurs even when the dominance and epistatic relations of genes (as reflected in relative values of genotypes) remain constant.

A. Evidence

A major portion of the evidence has been obtained in the form of estimates of genetic variances. Interpretation hinges closely on the nature of data used and the environmental base to which genotypic value is referred. An explanatory consideration of these issues will expedite both summary and discussion of the variance component estimates.

The major portion of variance estimates relate to the genetic differences among families or interactions of families with environment. It will suffice here to consider half-sib families; the prime source of evidence on additive genetic variance was data from replicated field comparisons of half-sib families. Four kinds of estimates have been obtained.
(i) Estimates of $\sigma_m^2$ from comparison of a set of families at more than one location in more than one year.

(ii) Estimates of $\lambda_1(\sigma_m^2 + \sigma_m^2 \lambda)$ from comparison of a set of families at one location in more than one year.

(iii) Estimates of $\lambda_1(\sigma_m^2 + \sigma_m^2 \mu)$ from comparison of a set of families in one year at more than one location.

(iv) Estimates of $\lambda_1(\sigma_m^2 + \sigma_m^2 \lambda + \sigma_m^2 \mu + \sigma_m^2 \mu \lambda)$ from comparison of a set of families at one location in one year.

$\sigma_m^2 = \text{genetic variance among half-sib families in terms of mean yield relative to a geographical region and a series of years (the geographical region visualized must be the one from which the locations of our experiments may be considered a random sample).}$

$\sigma_m^2 \lambda = \text{variance due to interaction of families with locations}$

$\sigma_{my}^2 = \text{variance due to interaction of families with years}$

$\sigma_{m\lambda y}^2 = \text{variance due to second order interaction of families, years and locations.}$

The four kinds of estimates can be viewed as differing in the amounts of bias involved.

Defining the value of a genotype in terms of average associated phenotype (yield) in all the environments of a region and series of years, $\lambda \sigma_m^2 = \text{the additive genetic variance plus a small (and specifiable) portion of the epistatic variance (see Part III, C).}$ The first of the four sorts of estimates is biased only by epistatic variance, the others by varying portions of the total genotype-environment interaction variance.

The differences among the kinds of estimates can also be viewed in terms of differential bases for the definition of genotypic value. Following in a general
way symbolism employed in Part III, let

\[ x = \text{value of a genotype defined as the average associated phenotype in all the environments of a region and a series of years} \]

\[ x_j = \text{value of a genotype similarly defined in terms of the environments at a specific location in a series of years} \]

\[ x_k = \text{value of a genotype in terms of the environments of a region in a specific year} \]

\[ x_{jk} = \text{value of a genotype in terms of the environments of a specific location in a specific year} \]

Let family means of \( x \), \( x_j \), \( x_k \) and \( x_{jk} \) be designated as \( m \), \( m_j \), \( m_k \) and \( m_{jk} \). Then

\[ 4(\sigma^2_m + \sigma^2_{m\ell}) = 4\sigma^2_{mj} \]

\[ 4(\sigma^2_m + \sigma^2_{my}) = 4\sigma^2_{mk} \]

and

\[ 4(\sigma^2_m + \sigma^2_{m\ell} + \sigma^2_{my} + \sigma^2_{my}) = 4\sigma^2_{mjk} \]

The estimate of each is an estimate of additive genetic variance (biased only by epistatic variance). The differences now reside in the way genotypic value is defined. For example, \( 4\sigma^2_m \) pertains directly to additive variance in genotypic value relative to the environments of a region and sequence of years, whereas \( 4\sigma^2_{mjk} \) pertains to additive variance in genotype value relative to a more homogeneous population of environments; those of a specific location in a specific year.

Finally, we should note that two sorts of evidence concerning genotype-environment interaction variance are available.
(i) The comparative sizes of the different sorts of estimates of genetic variance among families and
(ii) A number of direct estimates of $c^{2}_{m/y} + c^{2}_{m} + c^{2}_{m} + c^{2}_{m} + c^{2}_{m}.

1. Additive genetic variance

Two kinds of evidence have been obtained:
(i) estimates based on observed variation among random half-sib families and
(ii) the apparent effect of selection on the mean performance of populations.

Selection experiments have been conducted with three open-pollinated varieties and two hybrid populations. By hybrid population we mean the $F_2$ or any later generation (by cross pollination) from the cross of two pure lines. In all instances selection has been among full-sib families with mean grain yield as the sole criterion.

Direct estimates of additive genetic variance from observed family variation have been obtained for three open-pollinated varieties (the same three used in selection experiments), four hybrid populations (including the two used in selection experiments) and one variety cross population. By a variety cross population we mean the population derived with cross pollination from the $F_1$ of a variety cross. In the case of the variety cross population referred to above, the $F_1$ involved offspring from a large number of random matings between individuals of the two varieties.

The first, and fundamentally most important, aspect of the evidence is that existence of a moderate amount of additive variance is indicated. All
direct estimates have been positive and in four of the five selection programs
the apparent effect of selection has been consistent with the direct estimates
of additive variance in the populations concerned, (See Parts IV, C, l and IV, E).
The only evidence suggesting absence of additive variance is the lack of
established response to selection in one of the five selection experiments.

The best evidence concerning additive variance in \( x \), genotypic value
relative to the region and a sequence of years, has been obtained for the variety
Jarvis. It appears that the additive variance in the variety is in the range
from \( .0016 \) to \( .0024 \) where the unit of measurement is pounds of ear corn per
plant. The values, rather meaningless in themselves, can be made more tangible
in two ways. Considered in relation to estimates of total phenotypic variance
they indicate that heritability of grain yield in individual plants is between
5 and 8 percent. Considered with evidence concerning other components of the
phenotypic variance, they indicate enough additive genetic variance to allow 5
to 10 percent genetic improvement in mean population yield in response to one
cycle of selection among full-sib families. In arriving at these figures
(5 to 10 percent) no more than two replications in each of two years was
assumed as the basis for family comparisons. Instead of attempting to state
results for other populations in detail, we merely note that we have no
statistically significant evidence for population differences in additive genetic
variance and, with the exception of the variety cross population, no particular
reason to suspect major deviations from the situation in Jarvis. It would, of
course, be unrealistic to suppose that population differences do not exist;
they clearly must.
2. Relative magnitudes of dominance and additive genetic variance.

To minimize ambiguity in what follows we will begin by specifying the terms in which level or degree of dominance will be discussed. Consider any two alleles (B and b) in a random breeding population of infinite size in which genotype frequencies are at linkage equilibrium. Let

\[X_2\] be the mean genotypic value of the portion of the population that is BB at the locus of these alleles,

\[X_1\] be mean genotypic value of the portion that is Bb,

and \[X_0\] be mean genotypic value of the portion that is bb.

Unless \[X_2 = X_0\], one will be the larger. Assume \[X_2 > X_0\]. Then let

\[u = \frac{(X_2 - X_0)}{2}\]

and \[au = X_1 - \frac{(X_2 + X_0)}{2}\]

Then \[a\] becomes an index of the average dominance relation between B and b as follows:

(i) When \[X_1\] is exactly midway between \[X_0\] and \[X_2\] there is no dominance and \[a = 0\].

(ii) When \[X_1\] lies between \[X_0\] and \[X_2\] (but not exactly midway) there is partial dominance of one allele or the other and \[0 < |a| < 1.0\]. \[a\] will be positive if B is partially dominant, negative if b is partially dominant.

(iii) When \[X_1\] equals \[X_2\] or \[X_0\], there is complete dominance and \[a = 1\] if B is dominant \([X_1 = X_2]\) or \[a = -1\] if b is dominant \([X_1 = X_0]\).

(iv) When \[X_1\] lies outside the range \[X_0\] to \[X_2\], there is overdominance and \[|a| > 1.0\].

That dominance of some degree \((a \neq 0)\) exists in the action of genes bearing
on yield cannot be doubted. The related phenomena, inbreeding depression and heterosis in crosses of inbred lines, are almost universal and cannot be explained in terms of any model that does not allow $a > 0$ for one or more pairs of alleles. The necessary dominance may (or may not) be connected with some system of non-allelic gene interaction. Whether this is so (or not), the level of dominance is of interest and whether overdominance ($a > 1$) is a significant feature of gene action is of very special interest.

Our evidence is largely in the form of estimates of the relative sizes of dominance variance ($\sigma_d^2$) and additive genetic variance ($\sigma_g^2$) in the same eight populations for which estimates of additive genetic variance were obtained.

Two experimental designs (I and III) have been employed. Design III (used only with hybrid populations) provides information on whether estimates obtained with material from matings in the $F_2$ generation are biased upward by linkage. Results are presented in detail in Part IV, C. The most significant aspects were as follows:

(i) The estimates of $\sigma_d^2/\sigma_g^2$ were 1.17, 2.55, 1.83 and 1.51 for the four hybrid populations when $F_2$ generations were studied. The lowest value (1.51) was based on least data. Weighting in terms of amount of data the average estimate for hybrid populations was 1.50.

(ii) The estimates of $\sigma_d^2/\sigma_g^2$ were 1.17, 1.49 and 1.88 for the open-pollinated varieties. The weighted average for varieties was 1.43.

(iii) Design III trials for detection of linkage bias in estimates obtained for the $F_2$ of hybrid population have to date been
inconclusive. Decrease in $\sigma_d^2$ with generation advance from $F_2$ to $F_g$ was indicated in accord with expectation if $F_2$ estimates are biased as a result of linkage. However, there was a comparable trend in estimates of $\sigma_g^2$ so that the estimates of $\sigma_d^2/\sigma_g^2$ showed no certain trend.

It should be noted that the data from which individual estimates were obtained were all collected in a single year and location, in two years at one location, or at two locations in one year. Thus none pertain directly to $\bar{x}$, genotypic value defined with reference to the environments of a region and sequence of years.

Other results of some interest were as follows:

(i) Estimates of $\sigma_d^2/\sigma_g^2$ were rarely greater than .50, even in hybrid populations, for such characters as plant and ear height, tasseling date, ear number, ear diameter, ear length and kernel row number.

(ii) Estimates of $\sigma_d^2$ were negative (not statistically significant) for five of six characters studied in $N_j$ tabacum. The characters were days to flower, plant height, leaf number, leaf length, leaf breadth, and leaf yield. All exhibited highly significant amounts of additive genetic variance. The one positive estimate of $\sigma_d^2$ was only one-quarter as large as the estimate of $\sigma_g^2$ for the same trait.

3. Variance from interaction of genotype with environment

Interest centers around two issues. The first is the general magnitude of variance due to interaction of genotype with environment (as it varies with year and location) relative to variance in $\bar{x}$, genotypic value defined relative
to the environments of a region and a series of years. The second is whether the ratio of variance in $x$ to genotype environment interaction variance is the same for kinds of families that differ with respect to relative amounts of additive, dominance and epistatic variance contained in the total variance among families. Significance of this latter issue is discussed in Part III, A, l.

The available evidence pertaining to half-sib families in corn suggests that the total variance from interaction of families with years and locations, $\sigma^2_{m^2} + \sigma^2_{my} + \sigma^2_{my}$, is roughly equal or a bit greater than $\sigma^2_m$, the variance in family means of $x$. It appears that the three components of genotype environment interaction variance may be roughly equal in magnitude in the geographical region to which our data pertain. Data on yield in other crops (potatoes, soybeans, lespedeza and cotton) are in general agreement with the above for yield in corn. In the case of certain other characters most of which have probably been under less intense selection in the past, the genotype environment interaction variance is considerably smaller relative to genetic variance than in the case of yield.

The only evidence on the second issue is from corn data and it is not sufficient to be considered more than suggestive. At face value the evidence available indicates that, in the relative terms outlined above, the genotype environment interaction variance associated with non-additive genetic effects is smaller than that associated with additive genetic effects. In terms of symbolism employed in Parts II, A and III, A, this suggests that the relative values of additive dominance and epistatic variance may be different for $x$, $x_j$, $x_k$ and $x_{jk}$.

4. Epistasis

The evidence from corn data summarized in Part IV, D is by no means conclusive. However, at face value, it suggests that epistasis makes no very great contribution.
to genetic variance in the populations studied. Evidence obtained with other crops is presented in Part V, E. Preliminary results with cotton and soybeans, relative to the genetic variance among \( F_3 \) and \( F_4 \) families tracing by self-fertilization to different \( F_2 \) plants, are in general accord with the evidence from corn. On the other hand results with strawberries suggest that epistatic effects are of considerable magnitude in the material studied.

5. Genetic differences between open-pollinated varieties.

Heterosis in variety crosses has been estimated by comparison of six varieties with the fifteen possible crosses among them. Yield of each cross was greater than the average yield of its two parent varieties. The mean difference in yield between crosses and varieties was 19.9 percent of the mean yield of the varieties.

Estimates of the genetic variances among half-sib families and among full-sib families within half-sib families of the \( F_1 \) cross of Jarvis and Indian Chief have been obtained. In terms of comparable estimates (based on data from one location in one year) they were smaller than estimates of the same variances in the varieties themselves. The average estimate of genetic variance among half-sib families was 0.00049 for the cross and 0.00096 for the varieties. The average estimate of genetic variance among full-sib families within half-sib families was 0.00083 for the cross and 0.00126 for the varieties.

The same variances were estimated in one experiment using families obtained from matings of random plants from the \( F_2 \) of the Jarvis-Indian Chief cross. The estimate of genetic variance among half-sib families was 0.0021, the largest estimate of half-sib family variance to date in our work. However, the estimate for full-sib families within half-sib families was 0.00129 though this variance should be larger than the variance among half-sib families.
B. Genetic interpretations.

1. Additive genetic variance, dominance variance and genotype environment interaction variance.

The discussion of this section will proceed on the assumption that epistatic variance is an insignificant portion of the total genetic variance in populations of the sort studied. This is in line with the rather meagre evidence to date (see Part IV, E) and with a hypothesis about epistasis to be outlined in a later section.

Two facts appear certain. First, in the light of evidence concerning heterosis, there must be at least partial dominance of some, if not all, loci and this must be predominantly in the direction of the alleles with highest genotypic value when homozygous; symbolically $a > 0$ at the majority of loci. Second, in view of the evidence for genotype environment interaction, the relative genotypic effects (values) of genotypes at any given locus must be assumed variable with environment. Within these limitations it appears useful to distinguish three possible categories of loci.

(i) Loci at which one homozygote is superior to all heterozygotes and all other homozygotes (multiple alleles are visualized) in all admissible environments. The level of dominance involved is specified by saying $a < 1.0$ whether genotypic value is measured as $x$, $x_j$, $x_k$ or $x_{jk}$.

(ii) Loci at which one heterozygote is superior to any homozygote in most, if not all, of the admissible environments, i.e. $a > 1.0$ whether genotypic value is measured as $x$, $x_j$, $x_k$ or $x_{jk}$.

(iii) Loci at which homozygotes switch rank with variation in environment and one heterozygote, though usually inferior to one homozygote or
another in the environments of any single year-location, is on the average for all environments superior to all homoygotes. This means that the $u_i$'s are variable in sign as well as size with variation in environment and that $a > 1.0$ with respect to $x$, $a < 1.0$ with respect to $x_{jk}$ while $a$ may be either greater or less than 1.0 with respect to $x_j$ and $x_k$.

The evidence will be discussed in terms of each of these three kinds of loci. In doing so we must keep in mind that the amount of additive genetic variance and the ratio of dominance to additive variance are both dependent on gene frequencies. Gene frequencies in turn are dependent on past selection and the rate of introduction of new genes by mutation or migration. If the variation of environment is entirely random in time, selection causes gene frequencies to approach points of equilibrium which may be dominated solely by selection or jointly by selection and the rate of introduction of new genes. In discussion of varieties we shall consider two situations: gene frequencies at equilibrium and out of equilibrium. In all events, however, the frequency of the best allele (judged by genotypic value of the homozygote) will be assumed to be one-half or greater as a consequence of past selection. This is not a necessity for all loci but seems a reasonable assumption for the major portion of loci. Population geneticists are inclined to argue in terms of equilibrium gene frequencies and this may be generally proper for natural populations. However, in view of changes in corn cultural practices (especially fertilization) during the last 25 years it must be acknowledged that environmental variation has not been entirely random and hence that the average selective values of genes and therefore their frequency equilibrium points may have changed considerably. If so, it is easily possible that the processes
of selection have not had time in which to bring gene frequencies to their new equilibrium values. In hybrid populations at the \( F_2 \) level, gene frequencies at segregating loci will be assumed equal to one-half.

The first thing to be observed is that regardless of the type of locus that is predominant, gene frequencies are probably not at equilibrium in the open-pollinated varieties studied. Suggestive (but not conclusive) evidence is the apparent presence of additive genetic variance. More significant is the fact that response to selection in all three varieties was in agreement with prediction based on estimates of additive genetic variances with no allowance for any equilibrium forces. At equilibrium no response at all is expected unless selection for yield is made more intense than it had been in the past. While intensity of selection was probably increased (over prior selection in these varieties) in our selection experiments, response so close to prediction as actually observed would hardly occur if gene frequencies had been at equilibrium.

The hypothesis that all loci are of the first type is compatible with all our present evidence. Equilibrium gene frequencies in this case would be ones involving a balance between the forces of selection and mutation (or migration). These equilibria might at very high (greater than .95) or only moderately high (less than .95, even considerably less) frequencies of the best allele depending on rates of mutation and whether alleles most favorable to yield were unfavorable to other components of "fitness" such as germination, seedling viability, or production of effective pollen. In any event, even at equilibrium, the estimated additive genetic variance can be explained in terms of the segregation remaining if enough loci are assumed. Under this hypothesis,
the high estimates of $\sigma^2_d/\sigma^2_g$ in hybrid populations must be viewed as a reflection of pseudo-overdominance arising from linkage of partially dominant genes.

The hypothesis may be questioned on two counts: (i) the fact that Design III experiments for parallel estimation of $\sigma^2_d/\sigma^2_g$ in the $F_2$ and the $F_7$ or $F_8$ of hybrid populations have given no clear indication that linkage rather than true overdominance is responsible for high estimates at the $F_2$ stage, (ii) the acceptance by corn breeders by about 1930 that selection is ineffective for improvement of the average yield of adapted open-pollinated varieties. All that is worth saying now about the first issue is that data so far available are inconclusive. See Part IV for more detailed discussion of our own data. Limited evidence collected at the Nebraska Experiment Station (C. A. Gardner, personal communication) are favorable though not conclusively so, to the linkage bias explanation. Regarding the second issue, data from our own selection experiments are not in accord with the idea that selection in varieties is ineffective. This, of course, does not establish that selection would be equally effective in all other varieties. If there are varieties that cannot be improved by selection they would, under the hypothesis being considered, be explained as ones in which gene frequencies are at equilibrium and the argument presented by Crow (Genetics 33:478-487, 1948 and Heterosis, 282-297, 1952) then would be pertinent. He demonstrated that no cross of inbred lines extracted from the same "equilibrium" population is likely to be superior to the parent variety by more than 5 percent in mean yield. It is true that a superiority greater than 5 percent has sometimes been shown for such crosses, but it is not known whether the parent varieties were in those cases "equilibrium" populations. Some varieties (in particular the three used in our selection experiments) appear not to be. The general superiority
(10 to 20 percent) of selected inbred line crosses and double-crosses over varieties is not critical even if varieties are each at equilibrium so far as segregating loci are concerned. Selected crosses have usually been between inbred lines from different varieties and their excellence can be explained in terms of partial dominance on the grounds that certain good (and partially dominant) genes may be present in one variety but entirely lacking in another variety. This amounts to proposing that, relative to the whole specie, varieties should be viewed as partly inbred lines. The question of how varieties differ will be returned to in a later section.

The hypothesis that all loci are of the second type, $a > 1.0$ in all subsets of environments, can be rejected. If this were so, the equilibrium gene frequencies approached as a result of selection would be ones at which each locus contributed some dominance variance but no additive genetic variance. No convincing argument can be advanced for gene frequencies far enough from equilibrium to offer an explanation for the low estimates of $\sigma_d^2/\sigma_g^2$ from our data on varieties. However, it cannot be concluded that there are no loci of this type. It is possible that most of the dominance variance of varieties is associated with such loci and most of the additive genetic variance with loci at which there is low level partial dominance. This is not an attractive idea because one wonders why, if there is overdominance and low partial dominance, there is not also enough nearly complete dominance so that in total $\sigma_d^2/\sigma_g^2$ would be larger than appears to be the case. On the other hand, given a limited number of overdominance loci one might propose that the best single crosses of inbred lines are those heterozygous at many of these loci. In view of the fact that the hybrid populations we have worked with were initiated from high yielding single crosses it could be argued further that the high
estimates of $\frac{\sigma_d^2}{\sigma_s^2}$ obtained for these populations were a reflection of true
overdominance rather than of the joint effects of linkage and partial dominance.
If there were enough overdominant loci, linkage would still be a source of bias
in early generations, but as few as 10 or 20 such loci could be important in
terms of heterosis but distributed among the chromosomes in such a manner that
only one or two might be closely linked.

The proposal that there are some loci of each of the first two types
carries with it the suggestion that much of the superiority of selected inbred
line crosses over varieties is to be associated with overdominance loci
Crow, op.cit.). It is appropriate then to question whether this is possible
in view of the low estimates of dominance variance obtained for open-pollinated
varieties. The answer is yes. A further legitimate question is whether
response to selection (as noted in our experiments) would also be possible.
Again the answer is yes, but subject to two limitations: (i) only if gene
frequencies are not at equilibrium -- they need not be, and (ii) the limit of
varietal improvement through selection should fall short of the best attainable
cross of inbred lines from the same variety by something greater (perhaps
considerably greater) than 5 percent.

Finally, we need to consider the third type of locus; $a > 1.0$ for $x_j$
a $< 1.0$ for $x_{jk}$ and possibly for $x_j$ and $x_k$. The idea that the heterozygote
may be inferior to one or another homozygote in any specific environment (or
subset of environments) but superior to all homozygotes in the average for a
broad population of environments is one that has gained wide favor with popula-
tion geneticists in recent years. In ways pertinent to our evidence this sort
of locus is intermediate to the other two. Hence, if the evidence does not
exclude presence of some loci of each of the first two kinds it does not exclude loci of this third sort. Moreover, while the proportion of loci of the second sort that may be present is limited in certain ways by the evidence (particularly by evidence pertaining to \( \sigma_d^2 / \sigma_g^2 \) in varieties) there is presently not as severe a limitation on the number of this third kind. It represents a more subtle type of overdominance that can only be revealed or negated by data collected at more than one location in more than one year. Given only loci of this kind, pronounced genotype-environment interaction variance is to be expected and \( \sigma_d^2 / \sigma_g^2 \) would differ for \( x_i, x_j, x_k \) and \( x_{jk} \), being largest for \( x \) and smallest for \( x_{jk} \). Linkage bias in estimates of \( \sigma_d^2 / \sigma_g^2 \) in early generations of hybrid populations would be probable. Continuous response to selection would be expected in populations in which gene frequencies were at some distance from their equilibria but as equilibria were approached the apparent effect of selection might become variable in direction depending on temporal variations in environment. As with the second sort of loci, the limit of varietal improvement would fall short of the best cross between inbred lines from the variety. About all that can be said with much certainty concerning loci of this third (overdominant) sort is that presence of some is entirely plausible and they may in fact be of considerable significance.

2. Epistatic variance.

The evidence so far available suggests that epistatic variance comprises no large fraction of the total genetic variance of corn populations. In view of the almost inescapable evidence of many kinds that the effects of many genes having to do with fundamental physiological processes are interrelated, this may seem at first to be a highly improbable situation. On second thought, it is perhaps not so improbable. It may be that the genes which are most
important in terms of interrelated effects of other genes are the first to approach fixation as a consequence of selection in any random breeding population. In this connection, it is vital to note that the absence of epistatic variance in no way implies the absence of epistasis. If all segregating loci had a gene interaction relation with one or more homozygous locus but no such relations prevailed among the segregating loci, there would be a great deal of epistasis indeed but no epistatic variance. This idea has been examined in a preliminary way by mathematical analysis of the expected effects of selection in genetic systems involving gene interaction. The results are suggestive but not conclusive because the genetic models employed may have been too simple to reflect what happens in nature. The uniform finding was that selection eventually carried gene frequencies to sets of values such that epistatic variance is small relative to total genetic variance.

As a working hypothesis it is proposed that in any single random breeding population, with a long history of continuous selection by the same criterion and in the absence of drastic changes in environment, a core of genes that are most involved in non-allelic gene interactions will approach fixation so that epistatic variance becomes small relative to the rest of the residual genetic variance. It is further proposed that this "fixed genetic background" may vary from one population to another, an idea that appears inherent in Wright's "peaks of adaptation" (Amer. Nat. 90:5-24, 1956). It is suggested that the "genetic backgrounds" of two populations will tend to be similar or intrinsically different depending on whether the evolutionary separation of the populations was recent or remote on the evolutionary time scale. Presumably the similarity of "genetic background" of two populations would be reflected in the behaviour of material obtained by crossing the populations.
"F₂ breakdown", as observed in crosses of geographical races of *D. pseudo-
obscura* by Vetukhiv (Evolution 8:241-251, 1954), would be expected if the
"genetic backgrounds" were quite different. Likewise, the amount of epistatic
variance in the F₂ (or another early generation) of a population cross should
vary with the difference in "genetic backgrounds".

We have not noted "F₂ breakdown" in crosses of North Carolina varieties of
corn and have no information on epistatic variance in populations derived from
variety crosses.

A related proposal hinges on the fact that epistatic effects are associated
with unique combinations of genes at two or more loci and thus far segregating
loci have no permanence in cross fertilizing species that reproduce sexually.
Selection cannot preserve them (except by favoring fixation of a particular
"genetic background") against the recombination characteristic of sexual
reproduction by cross fertilization. On the other hand, unique and favorable
combinations can be preserved indefinitely in the pure lines of self-fertilizing
species or the clones of asexual specie. It has been proposed, therefore,
that significant amounts of epistatic variance are most likely to be found
in material from crossing pure lines of a self-fertilizing specie or the
clones of an asexual specie. Preliminary evidence has been in support of
this proposal in the case of strawberries but not in the case of cotton and
soybeans. The evidence from the latter two species lends credence to our
assumption in the previous section that epistatic variance is of secondary
consequence in the sort of corn populations from which most of our data has
been obtained.
3. Genetic differences between varieties.

The evidence that is most suggestive relative to genetic interpretations is on heterosis in variety crosses and on additive genetic variance in the $F_2$ from a variety cross. The considerable heterosis in variety crosses (20 percent difference in mean yield between 15 crosses compared with the six parent varieties) suggests that the varieties may be viewed as partly inbred lines, that one variety lacks certain useful genes that are present in one or more other varieties. Further evidence for this sort of difference between two particular varieties (Jarvis and Indian Chief) is provided from the $F_2$ of their cross where the one estimate of additive genetic variance was twice as great as the average of comparable estimates of additive genetic variance in the two varieties. This large increase of additive variance in the $F_2$ is most easily explained by postulating that more loci are segregating in the $F_2$, i.e., that at some loci the varieties were homozygous, or nearly so, for different alleles.

C. Summary in terms of significant objectives for future research.

All present evidence can be explained without recourse to assumption of overdominance in gene action relating to yield; at the same time presence of overdominance at some loci is not excluded. In this connection a limited number of results from other workers deserves attention. The results of Hull (op. cit.) from constant-parent regression analyses of data on inbred lines and their $F_1$ crosses are interpreted by him as indicative of overdominance. The validity of this interpretation is open to question, because the inbred lines involved were probably highly selected, but nevertheless must be kept in mind. On the other hand Sprague (in a paper read at the 1956 International Genetics Symposium, Tokyo) reported that when two populations were selected for combining ability with the same homozygous tester, the
result was increased yield in the cross of the two populations as well as in the
cross of each with the tester. Such a result is not expected if overdominance is
of prime importance. Various selection experiments with small animals
(e.g. Robertson and Reeve, J. of Genetics 50:141-148, 1952; Lerner and Dempster,
Heredity 5:75-94, 1951) have indicated effectiveness of selection for a time, then
ineffectiveness even though additive genetic variance of considerable magnitude
remained. The explanation appeared to lie in the fact that certain genes favorable
to the character under selection were unfavorable with respect to reproduction.

Further evidence relative to the overdominance issue is required. Some of
the key issues are as follows:

(i) Whether \( \sigma_d^2/\sigma_g^2 \) in varieties is larger for \( x \) than for \( x_{jk} \). This pertains
to the significance of loci at which a heterozygote is inferior to
one or another homozygote in the environments of a particular year-
location but superior to all homozygotes in the average for a region
and series of years. It requires data for estimation of \( \sigma_d^2 \) and \( \sigma_g^2 \)
from several locations and several years.

(ii) Whether high estimates of \( \sigma_d^2/\sigma_g^2 \) in early generations of hybrid
populations are due to linkage of partially dominant genes or to
true overdominance. Evidence can be obtained from continuation of
Design III experiments comparing estimates from the \( F_2 \) and advanced
generations of such populations.

(iii) The extent to which yield of a variety can be improved by selection.
The key issue is how close it can be brought to the yield of the best
cross of inbred lines from the same variety. An operational problem
is the population size required if adverse effects of inbreeding are
to be avoided.
Pertinent both to level of dominance in relation to the heterosis of corn hybrids and to differences among varieties are the following issues.

(iv) Whether additive genetic variance in populations obtained by crossing varieties is as high as indicated by preliminary evidence.

(v) The extent to which yield of populations obtained from the cross of two varieties can be increased by selection. Can it be made higher than what is possible from selection in the varieties themselves? Can it be made to approach the yield of the best crosses of inbred lines?

Very little experimental evidence on epistasis is now available. Much more is needed for evaluation of the proposals advanced on this subject. Evidence is needed for both varieties and populations obtained by crossing varieties. Experimental procedures are available but the amount of work involved is great.

D. Implications for corn breeding.

The most important aspects of present evidence with respect to corn breeding procedure are (i) the fact that neither partial dominance nor overdominance is yet excluded as a significant feature of the genetics of yield, (ii) the evidence for increased genetic variance in a population from a cross of two varieties and (iii) the fact that inbred line hybrids are to date so much superior to any other material that the focus of attention in corn breeding will certainly be on hybrids for years to come.

Reciprocal recurrent selection was proposed (Comstock, Robinson and Harvey, Agron. Jour. 41:360-367, 1949) as a breeding procedure likely to be effective regardless of level of dominance and superior to others if both partial and overdominance are important. In the light of present evidence it still appears a logical choice. Evidence on the performance of the $F_1$ cross between Jarvis and Indian Chief and on the genetic
variance among half-sib families in the $F_1$ of that cross indicates that reciprocal
selection using these varieties as the base populations might rather quickly result
in a variety cross equal or superior in yield to present North Carolina inbred line
hybrids. If so, the possibility is strong that still better performance might be
obtained in crosses of inbred lines extracted from the varieties as modified by
reciprocal selection.

A procedure that might be still better in the long run is suggested by the
indication of increased genetic variation in material from the cross of two varie-
ties. A double cross of four varieties should be available that has as high yield
as the Jarvis x Indian Chief cross. If so, a reciprocal selection program in which
the base populations were those derived from the two single crosses going into that
double cross would involve as high initial performance but greater potential for
ultimate improvement.
III. Theory

The objectives of the theoretical phase of our program have been outlined elsewhere (see Part I, 3). By their nature, activity in this portion of the work has been of two kinds. On the one hand, various distinct issues have called for rigorous mathematical analysis. Beyond this, however, wise planning and intelligent interpretation has demanded an ever broadening perspective through command of the genetic literature and mental synthesis of the evidence from both the literature and our own program. What follows is a condensation of our findings relative to specific issues. What has been accomplished otherwise will be reflected in the report as a whole and particularly Part II where we've attempted a condensation of the most important implications of present evidence and key issues for future study.

A. Clarification of issues associated with specification of the effect (or value) of a genotype

The variation of a qualitative character is variation in kind, e.g. flower color may be white, red, purple, etc. The effect of genotype on such traits is specified in terms of the kind of phenotype produced. On the other hand, the variation of quantitative characters is by definition, variation in quantity; and the specification of effect of genotype in terms of a metric of the phenotype is familiar and acceptable to the geneticist. (Here number of seeds, leaves, offspring, etc., is included as a special sort of metric.) However, there are associated problems that have never been clarified in the literature and to which we have of necessity given careful attention.

1. Choice of the base for evaluation of genotypes

Pertinent to the issue is the fact that phenotype (relative to quantitative aspects) is strongly affected by environment as well as by genotype and the fact that the totality of the environment of an individual is non-reproducible.
The idea that all the non-genetic variables bearing on the amount of grain produced by a corn plant could ever be both quantitatively and qualitatively the same for any two plants is almost inconceivable. This means that phenotype in a standard environment is unavailable as a practical measure of the values of genotypes. The practical alternative is to define genotypic value in terms of average phenotype relative to some infinite population of environments. The thing to note is that the relative values of genotypes may and doubtless do vary depending on the particular population of environments to which value refers. Hence consistent evidence concerning the values of genotypes or the variation of such values depends on constancy of the environmental base employed in the measurement of relative values and the choice of a particular base is required.

The nature of the choice for the plant breeder is almost automatic. By virtue of his objectives he is concerned with values of genotypes with respect to the environments associated with a geographical region during a sequence of years in the future. His only problem is one of setting boundaries in terms of space and admissible cultural practices.

There would be advantages to the geneticist in definition of genotypic value by reference to a more limited population of environments, in particular those genotypic values within the confines of a single location (experimental field) and a single year. If he does this, however, he will, in effect, be measuring value differently in successive years and, what is worse, in ways that allow limited inference to either the problems of the plant breeder or to the evolutionary history of his genetic materials.

We are led to the conviction that experimental data on the genetics of quantitative characters can have real utility only when it is referred to
values of genotypes based on the population of environments associated with a region and a sequence of years. This is appreciated in varying degrees by various other workers. Unfortunately, however, it is not so generally appreciated as to go without saying. A moderate proportion of the literature is written as if the authors (respected ones included) did not understand the problem or were content to leave their evidence applicable to only the restricted populations of environments in which their work was conducted.

2. Does evidence of genotypic diversity guarantee variation in genotypic values?

In general conformity we shall adopt for further purposes the following model for phenotype. It will be written with specific reference to the average phenotypic measurement for an experimental plot in a randomized block field trial but with minor modification in wording of the definitions of component effects would be equally applicable for phenotype of a single plant.

\[ P_{ijkl} = \mu_{jkl} + \bar{x}_i + (\bar{x}_l)_{ij} + (\bar{x}_y)_{ik} + (\bar{x}_l\bar{x}_y)_{ijk} + e_{ijkl} \]  

(1)

where \( \mu_{jkl} \) is the mean for the \( l \)-th block in the \( k \)-th year and \( j \)-th location

\( \bar{x}_i \) is the mean value of genotypes in the \( i \)-th family (the family may be of any kind and genetically either homogeneous or heterogeneous)

\( (\bar{x}_l)_{ij} \) is the average effect of interactions between genotypes of the \( i \)-th family and environments of the \( j \)-th location

\( (\bar{x}_y)_{ik} \) is the average effect of interaction between genotypes of the \( i \)-th family and environments of the \( k \)-th year

\( (\bar{x}_l\bar{x}_y)_{ijk} \) is the average effect of interactions between genotypes of the \( i \)-th family and environments of the \( j \)-th location and \( k \)-th year

\( e_{ijkl} \) is a residual not otherwise accounted for.

It is possible and useful to relate different measures of genotypic value in
terms of effects defined above. For example, the average genotypic value for the i-th family defined in terms of environments of a single year-location, say the j-th location and k-th year, is as follows:

$$\bar{x}_{ijk} = \bar{x}_i + (\bar{z}_l)_{ij} + (\bar{z}_y)_{ik} + (\bar{z}_{ly})_{ijk} \quad (2)$$

In like manner, the average genotypic value for the i-th family when defined in terms of environments of a sequence of years at a single location, say the i-th is as follows:

$$\bar{x}_{ij} = \bar{x}_i + (\bar{z}_l)_{ij} \quad (3)$$

The definition of the interaction effect $$(\bar{z}_l)_{ij}$$ is apparent from (3). It is the difference between average values of the genotypes of the i-th family defined on the one hand in terms of the environments of a region and a sequence of years, and on the other hand, in terms of the environments of a single location and a sequence of years.

In terms of either (2) or (3) it is now apparent that the answer to our initial question is No. For example, significant variation in the performance of families in a specific year-location is evidence only for variation in $$\bar{x}_{ijk}$$. This is evidence for genetic diversity because otherwise none of the four components of $$\bar{x}_{ijk}$$ could vary. On the other hand it does not guarantee variation in $$\bar{x}_i$$ since the variation in $$\bar{x}_{ijk}$$ could be entirely due to one or more of its other three components. To put this in slightly different words two or any number of families could be identical in value relative to a large population of environments but variable in genotypic value relative to any subset of the same population of environments. Complete absence of variance in $$\bar{x}_i$$ in the presence of genetic diversity among families seems improbable but the presence of considerable variance in $$\bar{x}_{ijk}$$ (or $$\bar{x}_{ij}$$ or $$\bar{x}_{ik}$$) can hardly be taken as a priori
evidence for large variance in $\bar{x}_1$.

This issue is vitally important in a practical sense because unbiased estimates of $\sigma^2_x$ (the variance of $\bar{x}$) require field comparison of families in more than one location and more than one year. This is costly but on the other hand there can hardly be an issue of more fundamental significance than the magnitude of $\sigma^2_x$.

3. Sources of variance in estimates of $\sigma^2_x$.

The genotypic variance among families in a genetic population is routinely estimated as a function of mean squares obtained in the analysis of variance of data from field comparison of a sample of families in two or more years and two or more locations. The procedure has frequently been described and need not be spelled out here. It is obvious that the vagaries of sampling will have a bearing on the actual estimate obtained. The sampling issue arises first in connection with families actually used and second in connection with the particular sample of environments (out of all those associated with the years and locations of the comparison) in which members of the different families happen to find themselves. The variance from this source (to be referred to hereafter as the "pure" sampling variance) may be considerable but it can be approximated reasonably well and brought within desired bounds by using sufficient families and enough replication.

Within the confines of the usual treatment of the subject of variance component estimation the source of variance in estimates of $\sigma^2_x$ considered in the preceding paragraph is the only source of such variance. This would be a very happy situation because it would mean that work in no more than two years and two locations could provide as much precision in the estimate as desired.
The matter appeared suspect and worthy of further examination. It turned out that the apparent (but questionable) simplicity of the situation revolved around assumption (1) that variances in genotype-environment interaction effects ($\sigma^2_{x\epsilon}$, $\sigma^2_{xy}$ and $\sigma^2_{x\epsilon y}$) are uniform from location to location or year to year, and (2) that all covariances among $\bar{x}$, ($\bar{x}\epsilon$), ($\bar{x}y$) and ($\bar{x}\epsilon y$) are zero in the years and locations of the comparison. Otherwise the sample of years and locations in which the comparison is conducted constitutes another source of variance in the estimate. This seems more reasonable than the earlier indication that what we've called "pure" sampling variance is the only source of error in estimates. However, it is operationally a very troublesome fact to face. It means that we cannot place final confidence in evidence concerning the size of $\sigma^2_x$ until estimates have been obtained in different samples of years and locations or until the variance from that source has been shown to be of minor consequence.

There appear to be no a priori grounds for arguing validity of the two assumptions listed in the preceding paragraph. Quite the contrary! This being so, the question reduces to one concerning the potential magnitude of errors associated with the sample of years and locations. The answer can only be provided by data. Once attention was focused on this question it was not difficult to indicate the type of data required. The requirement is met by data from comparison of a sample of families at three or more locations in each of three or more years. Given such data each variance component ($\sigma^2_x$, $\sigma^2_{x\epsilon}$, $\sigma^2_{xy}$, $\sigma^2_{x\epsilon y}$) can be estimated separately from the data associated with different sub-sets of the years and locations and the observed variation among estimates of a given variance component compared with the "pure" sampling variance expected. Data of the sort described is now being collected (see Part IV, C) and further work on the question is planned.
4. Estimation of the relative magnitudes of different components of the genetic variance

The comparative sizes of the additive genetic, dominance and epistatic variances of a genetic population constitute useful evidence relative to the nature of gene action. The amount of epistatic variance as a fraction of the total genetic variance provides an index to the importance of non-allelic gene interactions. The ratio of dominance to additive genetic variance is of special interest as evidence concerning level of dominance in gene action.

The relative sizes of the components of genetic variance are estimated in terms of the relative sizes of genetic variance among families of different kinds. For illustration it will suffice to consider one of the approaches (Design I) we have used in estimating the ratio of dominance variance ($\sigma_d^2$) to additive genetic variance ($\sigma_g^2$). Each of a series of random pollen parents is mated to a different set of four seed bearing parents. The result is a series of half-sib families each of which is made up of four full-sib families. The whole set of families is then compared in one or more field trials. Analysis of variance of the data supplies estimates of $\sigma_m^2$, the genetic variance among half-sib families, and of $\sigma_f^2$, the genetic variance among full-sib families within half-sib families. Assuming no epistasis and a population in which genotype frequencies are at linkage equilibrium,

$$\sigma_m^2 = \frac{1}{4} \sigma_g^2 \quad \text{and} \quad \sigma_f^2 = \frac{1}{4} \sigma_g^2 + \frac{1}{4} \sigma_d^2.$$ 

Then

$$\frac{\sigma_f^2}{\sigma_m^2} = \frac{\sigma_d^2}{\sigma_g^2}$$

and

$$\frac{\sigma_f^2}{\sigma_m^2} - 1 \quad \text{is an estimate of} \quad \frac{\sigma_d^2}{\sigma_g^2}.$$
It has been shown that when epistasis is present the amounts of epistatic variance contained in $\sigma_e^2$ is greater than the amount in $\sigma_m^2$ so that upward bias is introduced into the estimate of $\sigma_d^2/\sigma_g^2$ outlined above. However, the estimate is still useful in terms of a maximum bound for $\sigma_d^2/\sigma_g^2$.

An important procedural question arises in connection with these experiments. Will data from a single year and location suffice or must the comparison of the families be extended over two or more locations and two or more years? The question can be stated in another way using symbolism of a previous section. Consider the quantities $\bar{x}$ and $\bar{x}_{jk}$. The former is the average genotypic value for a family in terms of the environments of a region and a sequence of years. It is the measure of genotype about which we want information. The latter, $\bar{x}_{jk}$, is average genotypic value of a family in terms of the environments of a single year and location. Our question can be stated as follows: Is $\sigma_d^2/\sigma_g^2$ the same for $\bar{x}_{jk}$ as for $\bar{x}$? The information relative to $\bar{x}_{jk}$ can be obtained from the data of a comparison in a single year and location. While it cannot be considered obvious it is nevertheless true that the same precision of estimates requires a considerably smaller total quantity of data when the estimate pertain to $\bar{x}_{jk}$ rather than $\bar{x}$.

We once took the point of view that $\sigma_e^2/\sigma_m^2$ should not change by an important amount when the base for measurement of genotypic value is changed. Reconsideration has revealed that this point of view cannot be defended against all argument through it probably is correct for variances among two kinds of families to which the different kinds of genetic effects (additive, dominance and epistatic) contribute in the same proportions. For example, a hypothesis that has gained favor among population geneticists in recent years states that with respect to any segregating locus the heterozygote may be inferior to one
homozygote in many, if not most, individual environments but that it is less variable in value and for the average of a broad population of environments will, in many if not most instances, be superior to each homozygote. If this were the case it is not difficult to show that $\sigma^2_x/\sigma^2_m$ would be larger for $\bar{x}$ than for $\bar{x}_{jk}$.

From the foregoing it is abundantly apparent that the issue under discussion is even more important from a genetic than from an operational point of view. Preparations for obtaining experimental evidence on this question have been initiated.

B. Problems related to linkage

The problems investigated are all related to situations in which linkage causes deviation of genotype frequencies from those expected under independent segregation. This occurs when the number of genetically different gametes involved in formation of one or more generations of a population is small. The most extreme case is a population initiated from self-fertilization of a single plant or from a cross of two pure lines. In each case the population derives at one stage from only two gametes.

With the exception of some findings to be presented in Part III, C all our analyses of linkage effects have been made assuming no epistasis. Work previously reported dealt with the consequences of linkage relative to genetic variance component estimates obtained in Design I and Design III studies of hybrid populations (populations derived from the cross of two pure lines). The notable findings were that linkage offers a plausible explanation of the difference between varieties and hybrid populations with respect to estimates of $\sigma^2_d/\sigma^2_g$ and that Design III can be utilized to provide evidence relative to this explanation.
1. The effect of linkage on genotypic variances in populations obtained by self-fertilization after a cross of two pure lines

In cross-pollinating crops the passage of generations offers a continuing opportunity for recombinations among genes in the same chromosome. It is an obvious and widely appreciated fact that the probability of appearance of recombinant types is sharply reduced by initiation and continued use of self-fertilization. The effect of linkage in this situation has been rigorously specified in terms of genetic variances and covariances. The result obtained was in sufficiently general form to show the consequences of any number \( m \) of generations of random cross fertilization (beginning with the \( F_2 \)) prior to continuous self-fertilization. As indicated earlier absence of epistasis was assumed. The work was done as a part of a doctoral research problem.

Consider all the \( F_4 \) offspring (through self-fertilization) of an \( F_2 \) individual. Collectively they constitute what we call an \( F_2 \) subpopulation in the \( F_4 \) generation. In general an \( F_k \) subpopulation in the \( F_n \) generation consists of all \( F_n \) descendants (by selfing) of an individual \( F_k \) plant. If we specify that \( k' < k \), then any number of \( F_k \) subpopulations might trace to a single plant in the \( F_{k'} \) generation. A quantity of interest is the average genetic variance among \( F_k \) subpopulations tracing to a common ancestor in a generation (the \( F_{k'} \)) earlier than the \( F_k \) generation. The whole class of such variances may be symbolized as \((k, k'; n, n')\) where \( 1 \leq k' < k \leq n \). In the special case where \( n = k \) the average genetic variance among individuals within \( F_{k'} \) subpopulations is indicated. Another quantity of interest is the genetic covariance between subpopulation means in two generations \((n\) and \(n')\) for a set of \( F_k \) subpopulations tracing to a common ancestor in the \( F_{k'} \) generation. The whole set of such covariances is symbolized generally by \((k, k'; n, n')\). It happens that both
the variances and covariances indicated can be specified by a common formula.
Hence both variances and covariances will be symbolized by \((k, k'; n, n')\).
In the case of a variance \(n = n'\). It should be noted that when \(k' = 1\) the
\(F_{k'}\) generation is the \(F_1\). Within the \(F_1\) of a cross of pure lines there is only
one genotype, hence in effect there is only one possible common ancestor in the
\(F_1\) and only one \(F_1\) subpopulation which is the whole population derived from the
cross.

The expression that was derived is

\[
(k, k'; n, n') = \sum \left[ \frac{1}{2^{k'-1} - 1} - \frac{1}{2^{k-1}} + \frac{a_i^2 (2^k - 2^{k'})}{2^{n+n'-1}} \right] u_i^2
+ 4 \sum \sum \left[ \frac{q_{ij}^2 - r_{ij}^2}{1 - \pi_{ij}} + \frac{a_i a_j (q_{ij}^2 + r_{ij}^2)}{2^{n+n'-1}} \left( 4p_{ij} \right)^{k-2} - \left( 4p_{ij} \right)^{k' - 2} \right] u_i u_j
\]

Meanings of symbols are as follows:

\(u_i\) is half the difference in effect between the two homozygous genotypes
of the \(i\)-th locus.

\(a_i u_i\) is, for the \(i\)-th locus the difference between the effect of the
heterozygous genotype and the mean effect of the two homozygous
genotypes.

\(q_{ij}\) is the frequency of gametes, produced by the last generation from
cross pollination (the \(F_1\) when \(m = 1\)), of either coupling type with
respect to the \(i\)-th and \(j\)-th loci.

\(r_{ij}\) has similar meaning relative to gametes of repulsion type.

\[\pi_{ij} = \frac{1}{2} (1 - 2p_{ij})\] where \(p_{ij}\) is the recombination frequency between the
\(i\)-th and \(j\)-th loci.

\[\rho_{ij} = \frac{1}{2} (1 - 2p_{ij} + 2p_{ij}^2)\]
The number \( m \) of generations of cross-pollination between the \( F_2 \) and resumption of self-fertilization affects \( k, k', n, n' \) through \( q_{ij} \) and \( r_{ij} \). Assuming no selection in the course of cross pollination, \( q_{ij} \) and \( r_{ij} \) are functions of \( m \) and whether the linkage relation for the pair of loci concerned was coupling or repulsion in the \( F_1 \).

<table>
<thead>
<tr>
<th>State in ( F_1 )</th>
<th>Coupling</th>
<th>Repulsion</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q_{ij} )</td>
<td>( \frac{1}{2}(1 - p_{ij}) - D )</td>
<td>( \frac{p_{ij}}{2} + D )</td>
</tr>
<tr>
<td>( r_{ij} )</td>
<td>( \frac{p_{ij}}{2} + D )</td>
<td>( \frac{1}{2}(1 - p_{ij}) - D )</td>
</tr>
</tbody>
</table>

where \( D = \frac{1 - 2p_{ij}}{4} \left[ 1 - (1 - p_{ij})^m \right] \)

It is almost certainly an abstraction to talk of populations in which there is neither selection nor epistasis. Even so the general result given above provides a valuable point of departure for considering the effects of linkage. It has and will contribute to the problems of valid inference from estimates of variance and covariance of the sort considered. In addition it offers a guide to decisions concerning the value of cross-pollination (to enable more recombination) in normally self-fertilized crops. Implications of the general result have been given some attention from both of these points of view but will not be discussed here.

2. The consequences of linkage relative to the interpretation of mean phenotypic differences between heterozygotes and homozygotes in populations that are inbred.

Interest in overdominance as a possible source of persistent polymorphism in genetic populations has stimulated attempts to obtain evidence in various
ways. One approach has been to classify the individuals of a population according to genotype at a single locus (the phenotypic classes must be distinguishable by one means or another). The classes are then compared in terms of phenotypic means for one or more components of "fitness". A variant of this approach for which some data has been published involves comparison of the mean for heterozygotes with the highest of the means for homozygotes. A higher mean for heterozygotes was considered evidence for overdominance. This case was analyzed relative to possible effects of a linked locus. Let \( B \) and \( b \) represent the alleles at the locus on which classification is based. Let \( C \) and \( c \) be the alleles at a linked locus and let \( K \) be the contribution of the linked locus to the phenotypic mean difference between the BB and Bb classes (assuming BB class has a higher mean than the bb class). Then

\[
K = \frac{rs - qt}{(q + r)(s + t)} \left[ 1 - \left( \frac{q - r}{q + r} \right)^a \right] u
\]

where \( q = \) frequency of the BC gamete
\( r = \) frequency of the Bc gamete
\( s = \) frequency of the bC gamete
\( t = \) frequency of the bc gamete
\( u = \) half the difference in mean genetic effect between CC and cc
\( au = \) the difference between the mean genetic effect of Cc and the average of the mean genetic effects of CC and cc.

In populations with a long history of no inbreeding \((rs - qt)\) will be zero or very small so that \(K\) will also be zero or very small. The question now arises whether the expectation of \(K\) for inbred populations will be positive. There would appear to be equal probability of \((rs - qt)\) being positive or negative and hence some question whether on the average \(K\) would be positive. We have
not succeeded in obtaining a general expression for the expectation of $K$. However, we have obtained the result for populations established by one generation of full-sib mating and not inbred thereafter. The result was a positive value for $E(K)$. It must be qualitatively the same for inbreeding of any other kind or intensity.

The value of $E(K)$ may be small for any single locus, especially when inbreeding is mild. However every linked locus makes such a contribution and the total may be considerable without overdominance at any locus. This finding is not offered to discredit the sort of evidence under examination but only to indicate that due caution should be observed relative to inference from evidence of this kind.
C. Problems associated with epistasis

The interaction of non-allelic genes has engaged our attention in two ways: (1) as an aspect of the genetic variation of populations and (2) as a source of error in inference from theory which assumes absence of epistasis. The latter calls for brief comment. Mathematical treatment of various quantitative genetic problems has so far not been possible in terms of a completely general model. Results obtained (including some of great utility) have been produced under limiting assumptions. For example, assuming diploid meiosis, no linkage and no epistasis a great variety of problems are handled rather easily and many solutions have been provided in terms of this simple model. Allowing linkage but no epistasis does not ordinarily introduce insurmountable difficulties. Prior to 1954 no results of any generality relative to epistasis were available and even now only limited results are available pertaining to situations where linkage and epistasis are both present.

Thus grounds for judgement concerning the potential error in inferences which assume absence of epistasis have been and still are a pertinent object of investigation.

1. Errors involved in inference based on theory which assumes absence of epistasis.

This subject was explored in terms of four specific models for interaction among non-allelic genes assumed to segregate independently. Exhaustive results were obtained relative to: (1) the set of variances and covariances designated in Part III,B,1 as \(k, k'; n, n'\), (2) the variances of hybrid populations that have been investigated in the corn program and (3) the means for a pair of inbred lines and for generations obtainable from the pair of lines \(F_1\), any sort of backcross, \(F_2\) and \(F_n\) where the latter refers to any generation derived by continuous self-fertilization). These results were described in
our last comprehensive report and are published in detail in North Carolina Agricultural Experiment Station Technical Bulletin No. 118.

2. A generalized partitioning of the total genetic variance of a population.

The problem has been treated effectively by Cockerham who has worked within this group since 1953. He had made the steps that provided the key to his development of the subject before coming to North Carolina State College but has extended his results both in general and along lines pertinent to our research since joining this group. A parallel treatment of the primary aspects of the problem was provided by Kempthorne of Iowa State College at approximately the same time.

Fisher (op. cit) had defined three components of the total genetic variance: the additively genetic variance, the dominance variance and the remainder which he labeled epistatic variance. The key aspect of work by Cockerham and Kempthorne was a systematic (and meaningful) scheme for partitioning the total epistatic variance into a series of components identified with inter-locus interactions among the additive and dominance effects defined by Fisher. Any component in the set will be symbolized by $\sigma^2_{ad}$ where either $a$ or $d$ is allowed any value from 1 to $s$ (where $s$ is the number of loci affecting the character in question) subject to the restriction, $a + d \leq s$. Then $\sigma^2_{ad}$ is the total variance associated with interactions involving an additive effect at $a$ loci and a dominance effect at $d$ loci. There will be $(s - 1)$ of these components. The symbolism is normally extended to include additive genetic variance and dominance variance by letting $\sigma^2_{10}$ signify additive genetic variance and $\sigma^2_{01}$ signify dominance variance.

The partitioning of the epistatic variance was given utility by showing the fraction of each component contained in the genetic covariance between...
individuals related in different ways (parent and offspring, half-sibs, fullsibs, etc.). Both Cockerham and Kempthorne have done this for populations obtained by random mating of non-inbred parents. Two extensions were provided by Cockerham. The one deals with the genetic variances among families produced through continuous self-fertilization after a cross of two pure lines. A general result was provided for the class of variances identified earlier as \((k, k'; n, n')\). The other extension is concerned with genetic variances among and within families obtained from matings of inbred parents. The nature and significance of the results will be illustrated by attention to the special case when all parents are equally inbred and matings are between parents obtained randomly from the same base population. This case is directly relevant to proposed research using Designs I and II. Design I allows estimation of \(\sigma^2_m\), the genetic variance among half-sib families, and \(\sigma^2_f\), the genetic variance among full-sib families within half-sib families. Design II provides two estimates of \(\sigma^2_m\) and an estimate of \(\sigma^2_{mf}\), the genetic variance of full-sib families that is due to interaction of the genetic contributions of the two parents. Now let \(\sigma^2_s\) = the total genetic variance among full-sib families. Then

\[
\sigma^2_f = \sigma^2_s - \sigma^2_m
\]

\[
\sigma^2_{mf} = \sigma^2_s - 2\sigma^2_m
\]

The composition of \(\sigma^2_s\) and \(\sigma^2_m\) were found to be as follows:

\[
\sigma^2_s = \sum_{a,d} \left( \frac{1 + F}{2} \right)^a \left( \frac{1 + F}{2} \right)^{2d} \sigma^2_{ad}\tag{5}
\]

\[
\sigma^2_m = \sum_{a} \left( \frac{1 + F}{4} \right)^a \sigma^2_{a0}\tag{6}
\]
where $\sum_{a,d}$ signifies summations over all pairs of values of $a$ and $d$ subject to the restriction that $1 < a + d < s$, $\sum_{a}$ signifies summation over all values of $a$ from 1 to $s$ and $F$ is the inbreeding coefficient of the parents. Where $F = 1$, i.e. when the parents are pure lines extracted randomly from the population, $\sigma_s^2$ becomes $\sum_{a,d} \sigma_{ad}^2$ and is the total genetic variance. When $F < 1.0$, there will be a genetic variance within the full-sib families in the amount

$$\sigma_w^2 = \sum_{a,d} \left[ 1 - \left(\frac{1 + F}{2}\right)^a \left(\frac{1 + F}{2}\right)^{2d}\right] \sigma_{ad}^2$$

(7)

When $F = 0$, equations (5), (6) and (7) reduce to the forms applicable when non-inbred parents are employed.

Equations (5), (6) and (7) assume equal inbreeding of all parents and that the parents are an entirely random sample. A greater variety of variances can be estimated if the inbreeding of male parents is different from that of female parents, if parents belong to subsets such that within a single subset all parents trace to a common ancestor (the inbreeding being the same for all common ancestors or different for the ancestors of male parents from what it is for female parents), or if both of these are involved. The full result is known for any variation within this framework.

As indicated earlier, the foregoing results are all premised on independent segregation of interacting non-allelic genes.

3. Effects of epistasis in the presence of linkage.

The epistatic contributions to covariances among relatives have been analyzed for the special case of a random breeding population in which genotype frequencies are at equilibrium relative to linkage effects, i.e. the frequencies are those expected given independent segregation. In this situation linkage
has no effect when one relative is an ancestor of the other but does in the case of other kinds of relationships.

The consequences of linkage cannot be expressed in a general way or even very compactly for specific cases. For this reason results will be described only for the genetic covariance of half-sibs (this is equivalent to the genetic variance among half-sib families). When the parents are non-inbred and there is no linkage this is (see equation 6)

\[ \sigma_m^2 = \sum_a \left( \frac{1}{4} \right)^a \sigma_{a0}^2 \quad (a = 1, 2, \ldots, s) \]

Distinguishing the additive genetic variance from components of the epistatic variance

\[ \sigma_m^2 = \sigma_{10}^2 + \sum_a \left( \frac{1}{4} \right)^a \sigma_{a0}^2 \quad (a = 2, 3, \ldots, s) \]

In order to describe the linkage effects we must note that for any specific value of \( a \), \( \sigma_{a0}^2 \) contains a contribution from each possible set of \( a \) loci out of all those with genetic effects on the character. For each such set the linkage effect on the contribution of \( \sigma_{a0}^2 \) is potentially different. Now consider the adjacent pairs of loci within a set where adjacent implies (1) that the loci are located on the same chromosome and (2) that no other locus of the set lies between the two on the same chromosome. Subject to this limitation, there are a maximum of \( a - 1 \) adjacent pairs in any set of \( a \) loci. Assuming now that frequencies of multiple cross-overs are not affected by interference, the linkage effect is a function of the recombination frequencies for adjacent pairs of loci. Let

\[ p_i = \text{the recombination frequency for the } i\text{-th pair of adjacent loci, and} \]

\[ \delta_i = (1 - 2p_i^2) \]
Then the contribution of a given set of a loci to \( \sigma^2_{o0} \) is \( \eta (1 + \delta) \) times as large as if no linkage were present. Here \( \eta (1 + \delta) \) is the product of the values of \( (1 + \delta) \) for all adjacent pairs of loci and its value is limited to the range, zero to \( 2^{a-1} \).

In the case of other sorts of relatives (excluding those where one relative is an ancestor of the other and linkage has no effect), the linkage effect for any set of loci is also a function of the \( \delta \)'s but the value of the function depends on both the way individuals are related and whether the variance in question arises from interaction of additive effects, of dominance effects or of one or more additive effects with one or more dominance effects. Greater detail is given by Cockerham (Genetics 41:138-141, 1956).

It should be noted that results described above assume no position effects of genes.

It should also be noted that linkage effects in this case (which is very pertinent to our work because genotype frequencies must be close to linkage equilibrium in open-pollinated varieties) cannot be very great relative to total genetic variances unless evolutionary forces have been such that the loci of interacting genes tend to be located on the same chromosome instead of being scattered more or less randomly in the chromatin material. At the same time if linkages involved among an interacting set of genes are very close, any very immediate population effects of the set are as if linkage were complete so that for practical purposes the set behaves as one gene not involved in epistasis at all.
4. Design of experiments for obtaining information on the importance of epistasis.

Various approaches have been considered. Of these the three that appear to us of greatest utility are outlined below.

(a) The most promising of the approaches that can be employed in work with corn involves comparison of genetic variance estimates from Design I or Design II experiments in which the inbreeding coefficient of parental stocks is different. For example consider $4\sigma_f^2$ when $F$ (the inbreeding coefficient of parents) = 0 and $\sigma_s^2$ when $F = 1$ (each parent is a homozygous line).

When $F = 0$,

$$4\sigma_f^2 = 4 \left[ \sum_{a,d} \left( \frac{1}{2} \right)^{a+2d} \sigma_{ad}^2 - \sum_a \left( \frac{1}{4} \right)^a \sigma_{a0}^2 \right]$$

When $F = 1$

$$\sigma_s^2 = \sum_{a,d} \sigma_{ad}^2$$

The difference between these quantities involves only epistatic variance components as follows:

$$\sum_a \left[ 1 - \left( \frac{1}{2} \right)^{a-2} + \left( \frac{1}{4} \right)^{a-1} \right] \sigma_{a0}^2 + \sum_{a,d} \left[ 1 - \left( \frac{1}{2} \right)^{a+2d-2} \right] \sigma_{ad}^2$$

where $\sum_a$ and $\sum_{a,d}$ are used as before. Values of the bracketed coefficient of $\sigma_{a0}^2$ for various values of $a$ are as follows:

<table>
<thead>
<tr>
<th>$a$</th>
<th>$1 - \left( \frac{1}{2} \right)^{a-2} + \left( \frac{1}{4} \right)^{a-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>zero</td>
</tr>
<tr>
<td>2</td>
<td>$1/4$</td>
</tr>
<tr>
<td>3</td>
<td>$9/16$</td>
</tr>
<tr>
<td>4</td>
<td>$49/64$</td>
</tr>
<tr>
<td>5</td>
<td>$225/256$</td>
</tr>
</tbody>
</table>
The zero value when \( a = 1 \), indicates that \( \sigma_{10}^2 \), the additive genetic variance, does not contribute to the difference being considered. Values of the bracketed coefficient of \( \sigma_{ad}^2 \) for various values of \( a \) and \( d \) are as follows:

\[
\begin{array}{ccc}
\text{a} & \text{d} & \text{1 - (1/2) } a+2d-2 \\
0 & 1 & \text{zero} \\
0 & 2 & 3/4 \\
0 & 3 & 15/16 \\
1 & 1 & 1/2 \\
1 & 2 & 7/8 \\
2 & 1 & 3/4 \\
2 & 2 & 15/16 \\
\end{array}
\]

The zero value indicates that \( \sigma_{01}^2 \), the dominance variance does not contribute to the difference being considered. The coefficients for \( \sigma_{20}^2 \) and \( \sigma_{11}^2 \) are relatively small (1/4 and 1/2). Except for these two components the major portion of the epistatic variance is reflected in the difference being examined, and which can be estimated as outlined.

The contrast described above is not the only one that is possible from experiments of the sort indicated, but will serve for illustration.

(b) In work with asexual propagated crops \( \sigma_m^2 \) and \( \sigma_f^2 \) can be estimated using Design I or Design II and, because individual genotypes can be replicated, the genetic variance among full-sibs \( (\sigma_w^2) \) can be better estimated than is the case with corn. If data are obtained using half and full-sib families from non-inbreds parents the quantity, \( \sigma_w^2 - 3\sigma_f^2 + \sigma_m^2 \) will contain no additive or dominance variance but much of the epistatic variance. In terms of the epistatic components this quantity will be as follows:

\[
\sum_{a} \left[ 1 - \left( \frac{1}{2} \right) a-2 + \left( \frac{1}{4} \right) a-1 \right] \sigma_{e0}^2 + \sum_{a,d} \left[ 1 - \left( \frac{1}{2} \right) a+2d-2 \right] \sigma_{ad}^2
\]
This is the same quantity discussed in the preceding sub-section and will not be considered further.

(c) Designs I and II require extensive cross-pollination in preparation of the experimental material and hence are not convenient procedures with normally self-fertilized crops in which cross-pollination is difficult. For these we conceived the idea of obtaining information on the importance of epistasis by contrasting the parent-offspring covariances (or regressions) obtained with common parental stocks but two kinds of offspring. The procedure is outlined in Part V,E where preliminary results are reported. All that needs to be added here about the procedure is that a more powerful experiment than the one actually conducted might have resulted had the parental families been produced from individuals more highly homozygous than the $F_2$'s actually employed. The extreme in this connection, and possibly the best procedure, would involve use of pure lines derived randomly (or as nearly so as possible) from the population of interest.

The genetic difference between the two covariances has not been derived in terms of epistatic variance but the observed quantitative difference is important and directly pertinent in terms of one of the key questions in breeding work with self-fertilized crops. Considerable genetic variance is known to exist among families derived by selfing. Is the variance associated with additive gene effects of sufficient magnitude to yield continued improvement when a cyclic procedure of selecting among families and interbreeding selected families is followed?
D. The genetic difference among varieties

In corn, as in most other species, the dividing lines among sub-populations are in some places clear and sharp and in others indistinct and characterized by grades of intermediacy. Yet to understand a species it is necessary to understand the genetic interrelations, the differences and the similarities, among its sub-populations. The most tangible sub-populations of corn are the so-called open-pollinated varieties which once were the exclusive materials in corn production and now where displaced by hybrids are nevertheless the stocks from which hybrids were derived and which need to be understood if modern corn breeding is to be given a firm scientific basis.

There must be numerous sources of information (a portion doubtless unrealized to date) concerning the genetic interrelations of a pair of varieties. Among these certainly are (1) the relations among phenotypic means for the varieties, crosses between varieties, crosses of pure lines from a common variety, and crosses between pure lines from different varieties and (2) comparison of genetic variances within varieties and in populations derived in some way from the cross of two varieties. Some data relating to these issues have been collected and some attention has been given to theoretical bases for inference from such data (see Parts IV, D and III, C). We believe that the beginnings that have been made have tangible value but on the other hand, that they must be viewed as only beginnings which leave much that is vital to be explored in the future.

Any a priori sketch of the genetic issues that may be important relative to variety differences must view both multiple allalism and epistasis differently than considerations pertaining to the genetic variation within single populations. Nevertheless there is logic in delineating first the potential results of
Experimental data that could be explained without recourse to either multiple allelism or epistasis. For example, a set of experiments for which any results in the range possible would admit interpretation in terms of two alleles and no epistasis could obviously provide no information about multiple alleles and epistasis as factors in variety differences.

Expectations of genetic variances that can be estimated in experiments listed below have been derived for the two allele, no epistasis case.

1. Design I experiments in which families employed are obtained by mating random pollen parents from one variety with random seed parents from another variety.

2. Design I experiments in which families employed result from mating of random individuals from the F₁ or any later generation (derived with cross-pollination) of the cross of two varieties.

3. Design III experiments in which families employed are obtained by back-crossing random individuals of the F₁ or any later generation of a variety cross to both of the original parent varieties.

The results in all cases are expressed in terms of differences in effects between homozygotes at individual loci, level of dominance at the individual loci, and gene frequencies in the two varieties.

E. Reciprocal recurrent selection.

It is reasonable, on intuitive grounds, that the genetic improvement possible by this procedure should be a function of the genetic variance among the test cross progenies on which the selection is based. That this is so has been demonstrated and the form of the prediction equation based on estimates of this variance has been established.
One of the criticisms that has been leveled at reciprocal selection is that when gene frequencies at any overdominance locus are similar and close to the "equilibrium value" in the two populations, genetic change associated with that locus will be extremely slow. This is true but pertinent to this fact we have shown that whenever there is any difference between the populations in gene frequency at such loci, selection always operates to increase the difference and so to decrease the number of relatively unresponsive loci.

F. The variance of estimates (amounts of data required)

The utility of an experiment hinges on two things: whether it provides estimates of meaningful quantities and the amount of data required for adequate precision of those estimates. The latter issue has received attention relative to all experimental designs that have met the first requirement. Considered by itself this is statistical rather than genetical research, but it is entirely pertinent to effective use of resources and hence to productiveness of our program. No attempt will be made to summarize results except to note that they have had a bearing on all procedural decisions in our experimental program.
IV. Investigations with corn.

A. Introduction

Experimental information has been derived from three general sources:
(a) estimates of the components of phenotypic and genotypic variances (and co-
variances), (b) experiments for measuring effects of selection, and (c) experiments
for comparing means of populations of different sorts but genetically related, e.g. a
pair of inbred lines and the related $F_1$, $F_2$ and backcrosses. Maximum integration
of effort along these separate lines of attack has been an operating principle. For
example, variance studies are a part of all selection experiments in progress;
objectives include effects of selection on genetic variances as well as on means
and field work is designed so that the same data provides estimates of both means
and variances as well as the basis for selection of new parental stocks.

Our greatest interest has been in yield and only results pertaining to yield
will be fully reported. Results for other traits will be included only so far as
they seem specifically meaningful relative to principles of quantitative genetics.

Various experimental quantities and estimates have significance in more than
one context and for that reason will be recorded in more than one of the sections
that follow.

For reporting purposes it will be useful to list objectives in operational
terms as follows:

(1) Absolute magnitudes of (a) additive genetic variance and (b) the non-
genotypic components of phenotypic variance; variances due to interaction
of genotype and environment, and variance due to random effects of
environment.

(2) Relative magnitudes of the additive and dominance components of the total
genotypic variance.
(3) Effects of selection on means and genetic variances.
(4) Evidence relating to the presence and significance of epistasis (interaction of non-allelic genes).
(5) Evidence bearing on the nature of genetic differences between "open-pollinated varieties" of corn.
(6) Magnitude of genetic and non-genetic components of covariance between different quantitative characters.

Each of these immediate objectives has significance relative to both practical corn breeding and quantitative genetics. Beyond this, the ultimate objective is a pattern of experimental results that, together with information available from other sources, will enable a deeper insight concerning the quantitative genetics of corn populations.

This portion of the report deals separately with the topics listed above.

Part II, Discussion, is devoted to cross-meanings and joint interpretations.

B. Experimental material

The experimental program with corn is carried out on five farms operated by the North Carolina Agricultural Experiment Station and located in the Piedmont and Coastal Plain of the State. The amount of material in performance tests has been increased during the past five years from approximately 4500 plots in 1951 to almost 8000 test plots in 1956. The standard plot size used in all tests is two rows wide and ten hills in length with individual plants spaced 18 inches apart. The plots are planted at double rate of seeding and plants thinned to single plant hills when the plants are approximately six inches tall. Data are taken on ten equally competitive plants in each plot whenever possible, with special precautions taken to avoid non-random selection in all cultural operations in conducting the tests.
Data are taken on the following characters in most of the progeny tests: yield, date of tasseling, plant and ear height, ear number, ear length and ear diameter. Date of tasseling is recorded by plots at the time when approximately 50 percent of the plants are shedding pollen. Plant and ear height is measured in inches from ground level to the tip of tassel and uppermost ear bearing node, respectively. Ear length and ear diameter measurements are made only on the topmost ear of the plant and measured to the nearest one-tenth inch. Ear number and yield are based on all ears of the plants measured in a plot. Yield data are taken on unshelled ear corn at or near constant moisture content and may be referred to as "grain yield". Analyses of data are on plot values, obtained by dividing plot totals by the number of plants measured in each plot. Individual plant data on approximately every 12th plot are used in estimation of plant-to-plant variability within plots.

Seed stock nurseries are grown in the summer near Raleigh, North Carolina and in the winter near Homestead, Florida. A total of approximately 30,000 hand pollinations are made each year in the production of seed used in performance tests and in propagation of material used in the various populations. The winter nursery, enabling the production of two seed generations per crop year, has been used largely for rapid advance of inbred and random mated material. Future plans are to recombine selected parents of the selection studies in the winter nursery which will reduce the selection cycle by one year.

C. Magnitude of additive genetic variance, genotype-environment interaction variances and plot error variance

1. Grain yield
   a. Results

   All estimates of additive genetic variance ($\sigma_g^2$) have been based on estimates of the genetic variance among half-sib families ($\sigma_m^2$). Most
of the estimates have come from D-I experiments; a few from D-III experiments. In either case $\sigma^2_m$ contains no dominance variance. In D-I, $\sigma^2_m$ contains only small portions of the epistatic variance (maximum for any epistatic component is $\frac{1}{16}$ compared to $\frac{1}{4}$ of the additive genetic variance). The composition of $\sigma^2_m$ with reference to epistatic variance cannot be stated so specifically in the case of D-III, but at the present time it appears doubtful that there is any reason to be concerned over bias from epistasis in either D-I or D-III estimates (see Part III).

Estimates of $\sigma^2_g$ that are not biased by variance of effects due to interaction of genotype and environment require data from trials repeated in more than one year and more than one location. We have some data of each of the following sorts:

<table>
<thead>
<tr>
<th>Data from comparison of families in</th>
<th>Quantities estimable</th>
</tr>
</thead>
<tbody>
<tr>
<td>One location in one year</td>
<td>$\sigma^2_m + \sigma^2_{my} + \sigma^2_{ml} + \sigma^2_{my}$</td>
</tr>
<tr>
<td>One location in each of two years</td>
<td>$\sigma^2_m + \sigma^2_{my} + \sigma^2_{ml} + \sigma^2_{my}$</td>
</tr>
<tr>
<td>Two locations in one year</td>
<td>$\sigma^2_m + \sigma^2_{my} + \sigma^2_{ml} + \sigma^2_{my}$</td>
</tr>
<tr>
<td>Five locations in each of two years</td>
<td>$\sigma^2_m + \sigma^2_{my} + \sigma^2_{ml} + \sigma^2_{my}$</td>
</tr>
</tbody>
</table>

Estimates have been obtained for the following kinds of populations though much more extensively in some than in others.

1. Hybrid populations (derived from crossing two long time inbred lines), prior to other than natural selection.

2. Hybrid populations, after one or more cycles of family selection.

3. Open-pollinated varieties, as obtained.

4. Open-pollinated varieties, after one or more cycles of family selection in our program.
Data from a comparison of families in more than one year, location or both can always be divided into the portions obtained from the single year locations, and each of the portions used to provide a separate estimate of $\sigma_m^2 + \sigma_{my}^2 + \sigma_{ml}^2 + \sigma_{mly}^2$. This has been done in most instances because it enables more extensive comparison of the kinds of populations. It means, however, that some of the tabular material presented below is overlapping, i.e. evidence from the same data is presented in two different forms.

Table 4.1. Estimates from data collected in a single year and location ($\hat{\sigma}_g^2 = \frac{1}{4} \times$ estimate of $\sigma_m^2 + \sigma_{my}^2 + \sigma_{ml}^2 + \sigma_{mly}^2$).

<table>
<thead>
<tr>
<th>Kind of Population</th>
<th>Population</th>
<th>Design</th>
<th>Year</th>
<th>$\hat{\sigma}_g^2$</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid (unselected)</td>
<td>31 x 45</td>
<td>I</td>
<td>1947</td>
<td>.0019</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1955</td>
<td>.0060</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1956</td>
<td>.0018</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1956</td>
<td>.0027</td>
<td>.0031</td>
</tr>
<tr>
<td>CI21 x 7</td>
<td></td>
<td>I</td>
<td>1947</td>
<td>.0064</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1950</td>
<td>.0028</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1951</td>
<td>.0052</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1951</td>
<td>.0024</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1951</td>
<td>.0016</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1956</td>
<td>.0019</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1956</td>
<td>.0012</td>
<td>.0031</td>
</tr>
<tr>
<td>33 x K64</td>
<td></td>
<td>III</td>
<td>1950</td>
<td>.0040</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1951</td>
<td>.0040</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1951</td>
<td>.0032</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>III</td>
<td>1951</td>
<td>.0032</td>
<td>.0036</td>
</tr>
<tr>
<td>16 x 18</td>
<td></td>
<td>I</td>
<td>1947</td>
<td>.0070</td>
<td>.0070</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td>.0035</td>
<td></td>
</tr>
</tbody>
</table>

1/ All inbred lines used in the hybrids, except CI21 and K64, carry the NC prefix and were developed by Dr. P. H. Harvey of the North Carolina Experiment Station.
Table 4.1. (Cont'd)

<table>
<thead>
<tr>
<th>Kind of population</th>
<th>Population</th>
<th>Design</th>
<th>Year</th>
<th>$\hat{\sigma}_g^2$</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>O.P. variety (unselected)</td>
<td>Jarvis</td>
<td>I</td>
<td>1950</td>
<td>.0039</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1951</td>
<td></td>
<td>.0058</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1952</td>
<td></td>
<td>.0027</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1953</td>
<td></td>
<td>.0053</td>
<td>.0054</td>
</tr>
<tr>
<td></td>
<td>Weekley</td>
<td>I</td>
<td>1950</td>
<td>.0035</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1951</td>
<td></td>
<td>.0058</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1952</td>
<td></td>
<td>.0038</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1953</td>
<td></td>
<td>.0034</td>
<td>.0054</td>
</tr>
<tr>
<td></td>
<td>Indian Chief</td>
<td>I</td>
<td>1952</td>
<td>.0031</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1953</td>
<td></td>
<td>.0033</td>
<td>.0052</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td>.0042</td>
<td></td>
</tr>
<tr>
<td>Hybrid (selected)</td>
<td>31 x 45</td>
<td>I</td>
<td>1950</td>
<td>.0028</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1953</td>
<td></td>
<td>.0016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1956</td>
<td></td>
<td>.0002</td>
<td>.0015</td>
</tr>
<tr>
<td></td>
<td>21 x 7</td>
<td>I</td>
<td>1950</td>
<td>.0024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1951</td>
<td></td>
<td>.0016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1954</td>
<td></td>
<td>.0016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1955</td>
<td></td>
<td>.0026</td>
<td>.0020</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td>.0018</td>
<td>.0018</td>
</tr>
<tr>
<td>O.P. varieties (selected)</td>
<td>Jarvis</td>
<td>I</td>
<td>1954</td>
<td>.0002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1955</td>
<td></td>
<td>.0022</td>
<td>.0012</td>
</tr>
<tr>
<td></td>
<td>Weekley</td>
<td>I</td>
<td>1954</td>
<td>.0042</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1955</td>
<td></td>
<td>.0067</td>
<td>.0055</td>
</tr>
<tr>
<td></td>
<td>Indian Chief</td>
<td>I</td>
<td>1956</td>
<td>.0008</td>
<td>.0008</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td>.0028</td>
<td></td>
</tr>
</tbody>
</table>

**Grand averages**

Hybrid populations  .0030

O.P. varieties     .0037
Table 4.2. Estimates from comparisons at one location in each of two years
($\sigma^2_g = 4 \times \text{estimate of } \sigma^2_m + \sigma^2_{m\delta}$).

<table>
<thead>
<tr>
<th>Kind of Population</th>
<th>Population</th>
<th>Design</th>
<th>Years</th>
<th>Estimate of $\sigma^2_m + \sigma^2_{m\delta}$</th>
<th>$\sigma^2_{my} + \sigma^2_{m\delta y}$</th>
<th>$\sigma^2_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid (unselected)</td>
<td>21 x 7</td>
<td>III</td>
<td>50-51</td>
<td>.00030</td>
<td>.00030</td>
<td>.0012</td>
</tr>
<tr>
<td></td>
<td>33 x K64</td>
<td>III</td>
<td>50-51</td>
<td>.00025</td>
<td>.00070</td>
<td>.0010</td>
</tr>
<tr>
<td></td>
<td>Means</td>
<td></td>
<td></td>
<td>.00028</td>
<td>.00050</td>
<td>.0011</td>
</tr>
<tr>
<td>O.p. variety (unselected)</td>
<td>Jarvis</td>
<td>I</td>
<td>50-51</td>
<td>.00094</td>
<td>.00029</td>
<td>.0038</td>
</tr>
<tr>
<td></td>
<td>Jarvis</td>
<td>I</td>
<td>52-53</td>
<td>.00089</td>
<td>.00011</td>
<td>.0036</td>
</tr>
<tr>
<td></td>
<td>Weekley</td>
<td>I</td>
<td>50-51</td>
<td>.00031</td>
<td>.00098</td>
<td>.0012</td>
</tr>
<tr>
<td></td>
<td>Weekley</td>
<td>I</td>
<td>52-53</td>
<td>.00014</td>
<td>.00045</td>
<td>.0018</td>
</tr>
<tr>
<td></td>
<td>Indian Chief</td>
<td>I</td>
<td>52-53</td>
<td>.00035</td>
<td>.00044</td>
<td>.0014</td>
</tr>
<tr>
<td></td>
<td>Means</td>
<td></td>
<td></td>
<td>.00059</td>
<td>.00045</td>
<td>.0024</td>
</tr>
<tr>
<td>Hybrid (selected)</td>
<td>21 x 7</td>
<td>I</td>
<td>50-51</td>
<td>.00036</td>
<td>.00015</td>
<td>.0014</td>
</tr>
<tr>
<td></td>
<td>21 x 7</td>
<td>I</td>
<td>51-55</td>
<td>.00032</td>
<td>.00020</td>
<td>.0013</td>
</tr>
<tr>
<td></td>
<td>Means</td>
<td></td>
<td></td>
<td>.00034</td>
<td>.00016</td>
<td>.0014</td>
</tr>
<tr>
<td>O.p. variety (selected)</td>
<td>Jarvis</td>
<td>I</td>
<td>54-55</td>
<td>.00012</td>
<td>.00018</td>
<td>.0005</td>
</tr>
<tr>
<td></td>
<td>Weekley</td>
<td>I</td>
<td>54-55</td>
<td>.00046</td>
<td>.00051</td>
<td>.0034</td>
</tr>
<tr>
<td></td>
<td>Means</td>
<td></td>
<td></td>
<td>.00048</td>
<td>.00035</td>
<td>.0019</td>
</tr>
<tr>
<td>Grand averages</td>
<td>Hybrid populations</td>
<td></td>
<td></td>
<td>.00031</td>
<td>.00034</td>
<td>.0012</td>
</tr>
<tr>
<td></td>
<td>O.p. varieties</td>
<td></td>
<td></td>
<td>.00056</td>
<td>.00042</td>
<td>.0022</td>
</tr>
</tbody>
</table>
Table 4.3. Estimates from comparisons in two locations the same year
($\sigma^2_g = 4 \times \text{estimate of } \sigma^2_m + \sigma^2_{my}$).

<table>
<thead>
<tr>
<th>Kind of Population</th>
<th>Population</th>
<th>Design</th>
<th>Years</th>
<th>$\sigma^2_m + \sigma^2_{m\ell}$</th>
<th>$\sigma^2_{my} + \sigma^2_{m\text{ey}}$</th>
<th>$\sigma^2_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid (unselected)</td>
<td>21 x 7</td>
<td>III</td>
<td>1951</td>
<td>.00060</td>
<td>.00040</td>
<td>.0024</td>
</tr>
<tr>
<td></td>
<td>33 x K64</td>
<td>III</td>
<td>1951</td>
<td>.00040</td>
<td>.00060</td>
<td>.0016</td>
</tr>
<tr>
<td></td>
<td>Means</td>
<td></td>
<td></td>
<td>.00050</td>
<td>.00050</td>
<td>.0020</td>
</tr>
</tbody>
</table>

Table 4.4. Estimates from a comparison in five locations in each of two years
($\sigma^2_g = 4 \times \text{estimate of } \sigma^2_m$).

<table>
<thead>
<tr>
<th>Kind of Population</th>
<th>Population</th>
<th>Years</th>
<th>$\sigma^2_m$</th>
<th>$\sigma^2_{my}$</th>
<th>$\sigma^2_{m\ell}$</th>
<th>$\sigma^2_{m\text{ey}}$</th>
<th>$\sigma^2_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>O.p. variety (unselected)</td>
<td>Jarvis</td>
<td>55-56</td>
<td>.00065</td>
<td>-.00013</td>
<td>.00020</td>
<td>.00020</td>
<td>.0026</td>
</tr>
</tbody>
</table>

The average plot error variance for 25 experiments with over 200
degrees of freedom for error in each was .0058.

Estimates of the intra-plot variance among single plants have been
made regularly so that the total variance among plants could be estimated.
The average of 10 estimates of total variance in hybrid populations was
.0264. The average of 15 estimates of total variance in open-pollinated
varieties was .0336.

b. Discussion

(1) Reliability of estimates

Any very intelligent discussion of the size and variation among
the estimates that have been presented requires some notion of their
random variability. In Part III a distinction has been made between
(a) the "pure sampling" variance of these estimates and (b) that
associated with the location and years in which data were collected.
The former can be approximated reasonably well, the latter awaits
collection of empirical evidence (see Part III). To exemplify the
size of "pure sampling" variances, consider a genetic population and
sample of environments for which variances are of the approximate
magnitude computed on the average for open-pollinated varieties
(unselected). Most of the D-I estimates reported have involved compari-
sion of 64 half-sib families (each composed of 4 full-sib families) in
two replications at one location in one year (Table 4.1) or at one
location in two replications in each of two years (Table 4.2). For
such a population and such experiments, the "pure sampling" variances
of estimates would be as follows:

<table>
<thead>
<tr>
<th>Data from</th>
<th>Quantity estimated</th>
<th>Pure sampling variance</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>One year</td>
<td>(4(\sigma_m^2 + \sigma_{\text{ml}}^2 + \sigma_{\text{my}}^2 + \sigma_{\text{mly}}^2)) = .0043</td>
<td>.00000320</td>
<td>.0018</td>
</tr>
<tr>
<td>Two years</td>
<td>(4(\sigma_m^2 + \sigma_{\text{ml}}^2)) = .0025</td>
<td>.00000277</td>
<td>.0017</td>
</tr>
<tr>
<td></td>
<td>(\sigma_{\text{my}}^2 + \sigma_{\text{mly}}^2) = .00045</td>
<td>.00000064</td>
<td>.0008</td>
</tr>
</tbody>
</table>

Clearly, single estimates are subject to considerable random variation.
On the other hand the mean of several estimates acquires reasonably
small "pure sampling" variance. For example the mean of the seven
estimates of \((\sigma_m^2 + \sigma_{\text{ml}}^2)\) for open-pollinated varieties would have a
standard error from this source of \(.0017/\sqrt{7} = .00064\); about \(\frac{1}{4}\) of the
mean estimate itself.
Two points are worth noting: (a) That empirical checks on the "pure sampling" variance agree with values computed in the manner of those listed above. (b) That the observed variance among the seven estimates of $\sigma_m^2 + \sigma_{\bar{y}}^2$ for the open-pollinated varieties was actually somewhat smaller (.00000175) than the .00000277 given above.

(2) Differences between kinds of populations

The effect of selection on variances will be considered in greater detail in Part IV E. Here it will suffice to note (a) that results to date suggest some reduction in $\sigma_g^2$ as one consequence of selection and (b) that while one anticipates an eventual, if not immediate, decline in $\sigma_g^2$ as one of the general consequences of selection, it is far too early for even tentative conclusions about the rate of decline relative to time of duration of selection.

Mean differences observed between hybrid and varietal populations must be considered as no more than suggestive because of

(a) the small number of populations for which data have been collected.

(b) considerable lack of balance between the two kinds of populations with respect to years in which estimates were made.

Nevertheless, the means now available have some significance and are brought together in Table 4.5. The only conclusion is that present evidence establishes no striking differences (if indeed any at all) in additive genetic variance between the two kinds of populations.
Table 4.5 Average variance estimates for hybrid and open-pollinated variety populations.

<table>
<thead>
<tr>
<th>Variance estimate</th>
<th>Kind of Population</th>
<th>Hybrid</th>
<th>O.P. Variety</th>
<th>Ave.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_g^2 = 4(\text{estimate of } \sigma_m^2 + \sigma_{my}^2 + \sigma_{m\delta}^2 + \sigma_{m\delta y}^2)$</td>
<td>Hybrid O.P. Variety Ave.</td>
<td>.0030</td>
<td>.0037</td>
<td>.0033</td>
</tr>
<tr>
<td>$\sigma_g^2 = 4(\text{estimate of } \sigma_m^2 + \sigma_{m\delta})$</td>
<td>Hybrid O.P. Variety Ave.</td>
<td>.0012</td>
<td>.0022</td>
<td>.0018</td>
</tr>
<tr>
<td>$\sigma_m^2 + \sigma_{m\delta}$</td>
<td>Hybrid O.P. Variety Ave.</td>
<td>.00031</td>
<td>.00056</td>
<td>.00047</td>
</tr>
<tr>
<td>$\sigma_{my}^2 + \sigma_{m\delta y}^2$</td>
<td>Hybrid O.P. Variety Ave.</td>
<td>.00034</td>
<td>.00042</td>
<td>.00039</td>
</tr>
<tr>
<td>Plot error</td>
<td>Hybrid O.P. Variety Ave.</td>
<td>.0054</td>
<td>.0060</td>
<td>.0058</td>
</tr>
<tr>
<td>Total phenotypic</td>
<td>Hybrid O.P. Variety Ave.</td>
<td>.0264</td>
<td>.0336</td>
<td>.0312</td>
</tr>
</tbody>
</table>

(3) The approximate magnitude of additive genetic and other variances in open-pollinated varieties.

In this section, discussion will be phrased as if $\sigma_g^2$ and $4\sigma_m^2$ were exactly equal, i.e. the epistatic variance content of $\sigma_m^2$ and issues associated with genetic correlations between loci resulting from linkage, will be ignored. Limitation of discussion to varietal populations may be justified on the grounds that the varieties compose the basic genetic stocks from which hybrid populations are derived and to which therefore the latter owe their genetic character. More directly to the point perhaps is the fact -- that estimates for the two kinds of populations are about the same (see Table 4.5).

We have one unbiased estimate of $\sigma_g^2$ in an o.p. variety, the .0026 in Table 4.4. However, it was derived from only one variety and from data collected in only two years. Another source of information is the average estimate, .0024, of $4(\sigma_m^2 + \sigma_{m\delta}^2)$ in Table 4.2.
Possibly of the estimates in Table 4.4, that of $\sigma^2_{\text{m}^2}$ is more reliable than that of $\sigma^2_{\text{my}}$ because it is based on data from five locations whereas only two years were involved. At any rate it can be used to approximate the bias of the Table 4.2 estimate of $\sigma_g^2$. We obtain 

$0.0024 - 4(0.0002) = 0.0016$ as another unbiased estimate of $\sigma_g^2$ in o.p. varieties. Taking it all together, present evidence suggests that $\sigma_g^2$ is in or near the range, 0.0016 to 0.0026. So far as variances due to interaction of genotype with environment are concerned we have the following estimates:

<table>
<thead>
<tr>
<th>Variance</th>
<th>Estimate</th>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_{\text{my}} + \sigma^2_{\text{m}^2} y$</td>
<td>0.00142</td>
<td>4.2</td>
</tr>
<tr>
<td>$\sigma^2_{\text{my}}$</td>
<td>0.00013</td>
<td>4.4</td>
</tr>
<tr>
<td>$\sigma^2_{\text{m}^2}$</td>
<td>0.00020</td>
<td>4.4</td>
</tr>
<tr>
<td>$\sigma^2_{\text{m}^2} y$</td>
<td>0.00020</td>
<td>4.4</td>
</tr>
</tbody>
</table>

A variance cannot be less than zero and one feels strongly that there must be at least some genotype-year interaction, i.e. that $\sigma^2_{\text{my}} > 0$. This would be supported by the joint evidence of the estimates of $\sigma^2_{\text{my}} + \sigma^2_{\text{m}^2} y$ and of $\sigma^2_{\text{m}^2} y$ (see above). On the basis of evidence reviewed above it appears reasonable to propose that the sum of the three interaction variances ($\sigma^2_{\text{my}} + \sigma^2_{\text{m}^2} + \sigma^2_{\text{m}^2} y$) is as large or possibly a little larger than $\sigma^2_{\text{m}}$ and that the three may themselves not be greatly different in size.

Remembering that the total variance among plants averaged .0336 in o.p. varieties, the heritability ($\sigma^2_g / \sigma^2_p$) of grain yield in single plants would appear to be in or near the range $0.0016 / 0.0336 = 4.8$ to
.0026/.0336 = 7.7 percent.

2. Ear height and date of tasseling

Variance component estimates for these two characters are presented for contrast with results on yield. A smaller total of evidence is available than for yield but the contrast is so striking that only data on o.p. varieties (unselected) from comparisons in two years at one location will be given. For this group of populations estimates from the same studies are available for all three characters.

Table 4.6. Estimates from comparisons at one location in each of two years --
date of tasseling, o.p. varieties (unselected)

\[
(\delta^2_g = 4 \times \text{estimate of } \sigma^2_m + \sigma^2_{m'0})
\]

<table>
<thead>
<tr>
<th>Population</th>
<th>Years</th>
<th>(\sigma^2_{m+m'0})</th>
<th>(\sigma^2_{my+\bar{my}})</th>
<th>(\sigma^2_g)</th>
<th>(\sigma^2_e)</th>
<th>(\sigma^2_p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis</td>
<td>50-51</td>
<td>.71</td>
<td>.20</td>
<td>2.84</td>
<td>1.65</td>
<td>8.97</td>
</tr>
<tr>
<td>Jarvis</td>
<td>52-53</td>
<td>1.30</td>
<td>.07</td>
<td>5.20</td>
<td>1.48</td>
<td>10.79</td>
</tr>
<tr>
<td>Weekley</td>
<td>50-51</td>
<td>.84</td>
<td>.10</td>
<td>3.32</td>
<td>1.54</td>
<td>9.68</td>
</tr>
<tr>
<td>Weekley</td>
<td>52-53</td>
<td>1.96</td>
<td>.11</td>
<td>7.84</td>
<td>1.83</td>
<td>13.02</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>52-53</td>
<td>.98</td>
<td>.10</td>
<td>3.92</td>
<td>1.46</td>
<td>9.64</td>
</tr>
</tbody>
</table>

Means            |       | 1.16                  | .12                         | 4.62           | 1.59           | 10.42          |

* \(\sigma^2_e = \text{plot error variance, } \sigma^2_p = \text{total phenotypic variance among plants.}\)

Table 4.7. Estimates from comparisons at one location in each of two years --
ear height, o.p. varieties (unselected)

\[
(\delta^2_g = 4 \times \text{estimates of } \sigma^2_m + \sigma^2_{m'0})
\]

<table>
<thead>
<tr>
<th>Population</th>
<th>Years</th>
<th>(\sigma^2_{m+m'0})</th>
<th>(\sigma^2_{my+\bar{my}})</th>
<th>(\sigma^2_g)</th>
<th>(\sigma^2_e)</th>
<th>(\sigma^2_p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis</td>
<td>50-51</td>
<td>5.17</td>
<td>.60</td>
<td>20.68</td>
<td>6.41</td>
<td>44.83</td>
</tr>
<tr>
<td>Jarvis</td>
<td>52-53</td>
<td>2.47</td>
<td>.26</td>
<td>9.38</td>
<td>5.28</td>
<td>37.78</td>
</tr>
<tr>
<td>Weekley</td>
<td>50-51</td>
<td>6.74</td>
<td>.63</td>
<td>26.96</td>
<td>8.07</td>
<td>60.03</td>
</tr>
<tr>
<td>Weekley</td>
<td>52-53</td>
<td>6.70</td>
<td>.62</td>
<td>26.80</td>
<td>7.65</td>
<td>55.50</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>52-53</td>
<td>6.41</td>
<td>.48</td>
<td>25.64</td>
<td>5.55</td>
<td>39.71</td>
</tr>
</tbody>
</table>

Means            |       | 5.50                  | .52                         | 21.99          | 6.59           | 47.57          |

* \(\sigma^2_e = \text{plot error variance, } \sigma^2_p = \text{total phenotypic variance among plants.}\)
Table 4.8. Average variance estimates for three traits from comparisons at one location in each of three years. o.p. varieties (unselected)

<table>
<thead>
<tr>
<th>Variance</th>
<th>Yield</th>
<th>Ear height</th>
<th>Tasseling date</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_m$</td>
<td>.00059</td>
<td>5.50</td>
<td>1.16</td>
</tr>
<tr>
<td>$\sigma^2_{m\ell}$</td>
<td>.00045</td>
<td>.52</td>
<td>.12</td>
</tr>
<tr>
<td>$\sigma^2_y$</td>
<td>.00214</td>
<td>21.99</td>
<td>4.62</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>.0057</td>
<td>6.59</td>
<td>1.59</td>
</tr>
<tr>
<td>$\sigma^2_p$</td>
<td>.0328</td>
<td>47.57</td>
<td>10.42</td>
</tr>
</tbody>
</table>

Two points are worth special notice: (1) heritability is much higher for ear height and tasseling date than for yield, and (2) the apparent ratio of genetic variance to genotype-environment interaction variance (as judged from estimates of $\sigma^2_m + \sigma^2_{m\ell}$ and $\sigma^2_y + \sigma^2_{m\ell}$) is much less for yield than for the other two characters. These are emphasized by the following ratios of average estimates.

<table>
<thead>
<tr>
<th>Ratios of Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_g/\sigma^2_p$</td>
</tr>
<tr>
<td>$\sigma^2_m + \sigma^2_{m\ell}/\sigma^2_y + \sigma^2_{m\ell}$</td>
</tr>
</tbody>
</table>

3. Evidence on the sample of environments (locations and years) as a source of variation in estimates of additive genetic variance.

The theory of this issue is discussed in Part III, and types of evidence are outlined. Preliminary data of the required type have been obtained in a comparison of 60 half-sib families (from the Jarvis variety, unselected) at five separate locations in 1955 and 1956. The combined analysis of the data
yielded variance estimates reported in Table 4.4.

The data were further analyzed as outlined in Part III to provide evidence relative to (a) variation in family x environment interaction variances from year to year or location to location, or (b) non-zero covariances of family x environment interaction effects in different years or locations. The information to date is negative on both issues. However, more information is required before any very definite conclusions will be justified.
D. Comparative size of additive genetic ($\sigma_g^2$) and dominance ($\sigma_d^2$) variance.

Interest in the relative size of $\sigma_g^2$ and $\sigma_d^2$ relates to the information that the ratio supplies concerning the level of dominance exhibited by the genes affecting a character. The evidence to be summarized below was obtained in the same D-I and D-III experiments from which the variance estimates of Part IV-C were obtained. It has been pointed out in Part III that if upward biases in estimates of $\sigma_g^2$ and $\sigma_d^2$ are proportional to the respective genetic variances, then biased estimates are a better source of information on the ratio of the two genetic variances than unbiased estimates. Thus while data from a single year and location provide unsatisfactory information on the absolute size of the additive genetic variance, it may provide superior information on the ratio of $\sigma_d^2$ to $\sigma_g^2$.

1. Data relating to effect of bias in $\sigma_g^2$ and $\sigma_d^2$ on estimation of $\sigma_d^2/\sigma_g^2$.

Results from D-I studies repeated in two years are presented in Table 4.9. Similar data from D-III studies repeated in two years or two locations are presented in Tables 4.10 and 4.11. (See page 70 for these tables.) Two sorts of populations are involved:

(a) hybrid populations and

(b) c. p. varieties.

In the case of D-I data, $\sigma_m^2$ and $\sigma_f^2$ are considered nearly equal to $\frac{1}{4} \sigma_g^2$ and $\frac{1}{4} \sigma_g^2 + \frac{1}{4} \sigma_d^2$, respectively, so that $(\frac{\sigma_f^2}{\sigma_m^2} - 1)$ is the quantity pertinent to level of dominance. The issue of biased versus unbiased estimates resolves, therefore, into the question of whether $\sigma_f^2 + \sigma_{f\ell}^2$ and $\sigma_m^2 + \sigma_{m\ell}^2$ (quantities that can be estimated with two year data) and $\sigma_f^2 + \sigma_{f\ell}^2 + \sigma_{fy}^2 + \sigma_{f\ell y}^2$ and $\sigma_m^2 + \sigma_{m\ell}^2 + \sigma_{my}^2 + \sigma_{m\ell y}^2$ (the only quantities that can be estimated with single year, single location data) are of the same relative magnitude as $\sigma_f^2$ and $\sigma_m^2$. 
Table 4.9. Variance estimates from D-I studies repeated in two years at the same location.

<table>
<thead>
<tr>
<th>Population</th>
<th>Years</th>
<th>Estimates of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$A=\sigma_m^2 + \sigma_{m\ell}^2$</td>
</tr>
<tr>
<td>21 x 7</td>
<td>50-51</td>
<td>0.00036</td>
</tr>
<tr>
<td></td>
<td>54-55</td>
<td>0.00032</td>
</tr>
<tr>
<td>Jarvis</td>
<td>50-51</td>
<td>0.00094</td>
</tr>
<tr>
<td></td>
<td>52-53</td>
<td>0.00089</td>
</tr>
<tr>
<td></td>
<td>54-55</td>
<td>0.00012</td>
</tr>
<tr>
<td>Weekley</td>
<td>50-51</td>
<td>0.00031</td>
</tr>
<tr>
<td></td>
<td>52-53</td>
<td>0.00034</td>
</tr>
<tr>
<td></td>
<td>54-55</td>
<td>0.00086</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>52-53</td>
<td>0.00035</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td>0.00051</td>
</tr>
</tbody>
</table>

Table 4.10. Variance estimates from D-III studies repeated in two years at one location.

<table>
<thead>
<tr>
<th>Population</th>
<th>Years</th>
<th>Estimates of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$A=\sigma_m^2 + \sigma_{m\ell}^2$</td>
</tr>
<tr>
<td>21 x 7</td>
<td>50-51</td>
<td>0.00030</td>
</tr>
<tr>
<td>33 x K64</td>
<td>50-51</td>
<td>0.00020</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td>0.00025</td>
</tr>
</tbody>
</table>

Table 4.11. Variance estimates from D-III studies repeated at two locations in one year.

<table>
<thead>
<tr>
<th>Population</th>
<th>Years</th>
<th>Estimates of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$A=\sigma_m^2 + \sigma_{m\ell}^2$</td>
</tr>
<tr>
<td>21 x 7</td>
<td>51</td>
<td>0.00060</td>
</tr>
<tr>
<td>33 x K64</td>
<td>51</td>
<td>0.00040</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td>0.00050</td>
</tr>
</tbody>
</table>
If not, it seems reasonable to assume that the least biased quantities would be closer in relative size to $\sigma_f^2$ and $\sigma_m^2$.

In the case of D-III data the quantity to be estimated is $\sigma_d^2 / 4 \sigma_m^2$ and the same sort of questions arise concerning the biased estimates of $\sigma_d^2$ and $\sigma_m^2$ available from two-year or two-location data and from single year-location data.

These considerations suggest a comparison of the ratios $\hat{B}/\hat{A}$ and $\hat{D}/\hat{C}$. Employing average values of the estimates of $A$, $B$, $C$ and $D$ the following values are obtained:

<table>
<thead>
<tr>
<th></th>
<th>$\hat{B}/\hat{A}$</th>
<th>$\hat{D}/\hat{C}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-I data</td>
<td>.73</td>
<td>.66</td>
</tr>
<tr>
<td>D-III data</td>
<td>.73</td>
<td>.16</td>
</tr>
</tbody>
</table>

One could compute the ratios separately for each set of estimates and then average, but such averages are badly biased (upward) when the denominator is subject to a great deal of random error as in these cases. It is also quite certain the bias would be different for the two ratios.

The above figures indicate the possibility that when biased estimates are employed the bias may be somewhat greater in estimates of $\sigma_g^2$ than in those of $\sigma_d^2$. Put differently, unbiased estimates (which are very costly to obtain with accuracy) might indicate higher levels of dominance. While the data are variable and far from conclusive, the issue is so important that whatever evidence can be found is worth examining. In this connection it should be noted that the genetic variance involved in $D$ is only partly non-additive in the case of Design I but entirely non-additive in the case of Design III. Thus the contrast between the D-I and D-III results is in line with the possibility that the ratio of genotype-environmental interaction variance to genetic variance is lower for non-additive than for additive genetic effects.
2. Estimates of $\sigma_d^2$ and $\sigma_g^2$ in o.p. varieties and the F$_2$ generation of hybrid populations.

   a. Grain yield

   Table 4.12 lists values of $\sigma_d^2$ and $\sigma_g^2$ obtained for hybrid populations studied at the F$_2$ stage. Table 4.13 contains the same information for o.p. varieties.

   Table 4.12. Estimates of $\sigma_d^2$ and $\sigma_g^2$

<table>
<thead>
<tr>
<th>Hybrid Populations, F$_2$ Parental Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>21 x 7</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>34 x 45</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>33 x k34</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>16 x 18</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* Nos. indicate design, number of years, and number of locations. Thus, III-2-1 indicates Design III in two years at one location.
Table 4.13. Estimates of $\sigma_d^2$ and $\sigma_g^2$

<table>
<thead>
<tr>
<th>Population</th>
<th>Design</th>
<th>Years</th>
<th>$\sigma_d^2$</th>
<th>$\sigma_g^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis</td>
<td>I-2-1</td>
<td>50-51</td>
<td>0.0043</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>I-2-1</td>
<td>52-53</td>
<td>0.0038</td>
<td>-0.0012</td>
</tr>
<tr>
<td></td>
<td>I-2-1</td>
<td>54-55</td>
<td>0.0008</td>
<td>0.0020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Means</td>
</tr>
<tr>
<td>Weekley</td>
<td>I-2-1</td>
<td>50-51</td>
<td>0.0032</td>
<td>0.0026</td>
</tr>
<tr>
<td></td>
<td>I-2-1</td>
<td>52-53</td>
<td>0.0027</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>I-2-1</td>
<td>54-55</td>
<td>0.0045</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Means</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>I-2-1</td>
<td>52-53</td>
<td>0.0023</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>I-1-1</td>
<td>56</td>
<td>0.0008</td>
<td>0.0020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Means</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grand Means</td>
</tr>
</tbody>
</table>

All listed values of $\sigma_d^2/\sigma_g^2$ are computed from means of $\sigma_d^2$ and $\sigma_g^2$. For example, in table 4.12, $1.17 = 0.0034/0.0029$ and $1.50 = 0.00486/0.00324$. The reason for so proceeding instead of computing $\sigma_d^2/\sigma_g^2$ for each experiment and then taking averages of the individual values of $\sigma_d^2/\sigma_g^2$ is that averages of this latter type are more seriously biased (upward).

The two significant issues in the data of Tables 4.12 and 4.13 are

(a) $\sigma_d^2/\sigma_g^2$ is over three times as large, in the average, for hybrid populations as it is for o.p. varieties. It is larger for every hybrid population than for any variety, though the lowest value for hybrids and the highest for varieties are of the same general size.

(b) Values of $\sigma_d^2/\sigma_g^2$ for hybrid populations are uniformly in excess of
50. The significance of this is that for the variance arising from a single locus $\sigma_d^2/\sigma_g^2$ is greater than .5 only in the event of overdominance. Thus large values of $\hat{\sigma_d^2}/\sigma_g^2$ suggest presence of overdominance generally and must be so interpreted in absence of a tenable alternative.

There is moderate variation in $\sigma_d^2/\sigma_g^2$ among the hybrid populations. This is probably not statistically significant. On the other hand there is little doubt that real differences among populations exist so the issue of statistically significant differences is beside the point.

b. Other characters.

Comparative estimates of $\sigma_g^2$ and $\sigma_d^2$ will not be presented for characters other than yield. However, they have been made and recorded for all characters that have been measured in our experiments (See Part IV, B). The whole matter can be summarized by stating that $\hat{\sigma_d^2}/\sigma_g^2$ has only rarely been greater than one-half for traits other than yield, and that this is true for hybrid as well as varietal populations. Thus, the weight of the evidence is that in these traits dominance is less of a factor than for grain yield and that for the most part dominance is probably less than complete.

3. Evidence relative to linkage effects on $\sigma_d^2/\sigma_g^2$ in hybrid populations.

In terms of the entire genotype linkage prevents independent segregation of many pairs of loci. Nevertheless in populations in which mating has been approximately random for many generations, genotype frequencies will be nearly
as they would be if segregation were actually random. This, however, will not be the case in the \( F_2 \) generation of a hybrid population and will be approached only slowly and asymptotically in generation time. This consequence of linkage provides a possible explanation (alternative to overdominance) of the large values of \( \sigma_d^2/\sigma_g^2 \) for hybrid populations listed in Table 4.13.

If linkage is an important issue relative to findings in the hybrid populations, the situation should change as opportunity is provided for recombination to dissipate the effects of linkage on relative frequencies of genotype. Evidence of two sorts has been accumulated.

Family selection has been practiced in two hybrid populations, \( 21 \times 7 \) and \( 34 \times 45 \). During every cycle of selection new estimates of \( \sigma_g^2 \) and \( \sigma_d^2 \) are obtained (in D-I experiments) for comparison with original estimates. The successive estimates for these populations are listed in Table 4.14. Successive estimates are separated by a two generation span in every instance. If linkage was a factor contributing to high values of \( \sigma_d^2/\sigma_g^2 \) in experiments based on \( F_2 \) parental material, the successive values should show a downward trend. A similar trend should occur in \( \sigma_d^2 \) but the behaviour of \( \sigma_g^2 \) cannot be forecast in terms of response to recombination.

<table>
<thead>
<tr>
<th>Table 4.14. Successive estimates of ( \sigma_g^2 ) and ( \sigma_d^2 ) in two hybrid populations.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>34 x 45</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>21 x 7</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Table 4.15 contains results from D-III studies with the same populations. In the case of 34 x 45 the \( F_2 \) was contrasted experimentally with the \( F_6 \) from the selection experiment with this population. In the 21 x 7 population the \( F_2 \) was contrasted with an \( F_8 \) developed with minimum selection.

**Table 4.15.** Design III estimates of \( \sigma_g^2 \) and \( \sigma_d^2 \) in two hybrid populations at different generations.

<table>
<thead>
<tr>
<th>Population</th>
<th>Year</th>
<th>( \sigma_g^2 )</th>
<th>( \sigma_d^2 )</th>
<th>( \sigma_d^2 / \sigma_g^2 )</th>
<th>( \sigma_g^2 )</th>
<th>( \sigma_d^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>34 x 45</td>
<td>55</td>
<td>.0060</td>
<td>.0057</td>
<td>.00323</td>
<td>.0012</td>
<td>.0021</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>.0018</td>
<td>.0042</td>
<td>.00403</td>
<td>.0011</td>
<td>.0013</td>
</tr>
<tr>
<td>21 x 7</td>
<td>56</td>
<td>.0019</td>
<td>.0022</td>
<td>.00117</td>
<td>.0012</td>
<td>.0014</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td>.00323</td>
<td>.00403</td>
<td>.00114</td>
<td>.00117</td>
<td>.00160</td>
</tr>
</tbody>
</table>

The ratio \( \sigma_d^2 / \sigma_g^2 \) is as large in Table 4.15 for advanced generations as it is for the \( F_2 \). In Table 4.14 the ratio varies greatly and exhibits no clear trend. Actually the observed ratio is subject to large sampling variance and for this reason provides less reliable information than the separate estimates of \( \sigma_d^2 \) and \( \sigma_g^2 \). In view of the considerable sampling variance of even these estimates it seems worthwhile to average all available estimates (Tables 4.12, 4.14 and 4.15). This means putting together results from selected and unselected advanced generation material in one instance (the \( F_6 \) generation of 34 x 45). Averaging is otherwise confined to the \( F_2 \) where selection was never involved.

Results are listed in Table 4.16. (See p. 77) The impression gained from these figures, or from Tables 4.14 and 4.15, is that \( \sigma_d^2 \) is decreasing with generation advance but that a roughly comparable decrease was taking place in \( \sigma_g^2 \). Taking the results at face value, a decrease in \( \sigma_d^2 \) is consistent with the
Table 4.16. Estimates of \( \sigma_g^2 \) and \( \sigma_d^2 \) by generations in two hybrid populations.

<table>
<thead>
<tr>
<th>Population</th>
<th>Generation</th>
<th>No. of Estimates</th>
<th>( \hat{\sigma}_g^2 )</th>
<th>( \hat{\sigma}_d^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>34 x 45</td>
<td>( F_2 )</td>
<td>4</td>
<td>0.0037</td>
<td>0.0079</td>
</tr>
<tr>
<td></td>
<td>( F_4 )</td>
<td>1</td>
<td>0.0028</td>
<td>0.0068</td>
</tr>
<tr>
<td></td>
<td>( F_6 )</td>
<td>3</td>
<td>0.0013</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>( F_8 )</td>
<td>1</td>
<td>0.0002</td>
<td>0.0013</td>
</tr>
<tr>
<td>21 x 7</td>
<td>( F_2 )</td>
<td>4</td>
<td>0.0029</td>
<td>0.0034</td>
</tr>
<tr>
<td></td>
<td>( F_4 )</td>
<td>1</td>
<td>0.0020</td>
<td>0.0056</td>
</tr>
<tr>
<td></td>
<td>( F_6 )</td>
<td>1</td>
<td>0.0021</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>( F_8 )</td>
<td>1</td>
<td>0.0012</td>
<td>0.0014</td>
</tr>
</tbody>
</table>

The notion of a linkage effect associated with a multiplicity of loci. A comparable trend in \( \sigma_g^2 \) suggests a preponderance of coupling in the \( F_2 \) generation. It should be reiterated here that the theoretical effect of recombination (break-up of \( F_2 \) linkage relations) is a downward trend in \( \sigma_d^2 \) that is independent of the nature of the original linkage, but is an upward, downward or no trend in \( \sigma_g^2 \) depending on the relative amounts of repulsion and coupling in the \( F_2 \). However, \( \sigma_g^2 \) should in no circumstance decrease more rapidly than \( \sigma_d^2 \). A pertinent question at this time is whether the evidence is sufficient to establish relative rates of change in the two variance components in a sufficiently reliable way. The answer is No.

The general impression left by Table 4.16 is highly colored by the \( F_8 \) estimates. If \( \hat{\sigma}_g^2 \) had been somewhat higher and \( \hat{\sigma}_d^2 \) somewhat lower for that generation tentative indications would be rather different. Hence for illustration, it will serve to consider the reliability of the single estimates obtained for that generation. Further it will suffice to consider the sampling errors of these estimates assuming as indicated by evidence to date (see Parts
IV C-1-a and IV C-3) that the sample of environments in which estimates are obtained is not a major source of variation among estimates.

Let us examine the estimates for the 21 x 7 population which, by reason of the way they were obtained and amounts of data, are subject to less sampling variance than the $F_8$ estimates for 34 x 45. Suppose the true variances had been

$$
\sigma_g^2 = .0024 \quad \text{and} \quad \sigma_d^2 = .0010.
$$

These in place of the .0012 and .0014 of Table 4.16 would have suggested no trend in $\sigma_g^2$ and a major decrease in $\sigma_d^2$. Given $\sigma_g^2 = .0024$ and $\sigma_d^2 = .0005$ and plot error variance as observed in the experiment the minimum standard errors of the estimates would have been as follows.

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_g^2$</td>
<td>.00072</td>
</tr>
<tr>
<td>$\hat{\sigma}_d^2$</td>
<td>.00033</td>
</tr>
</tbody>
</table>

It is quite clear that the estimates obtained may be sampling deviates from true values that, if known, would place the whole matter in quite different perspective.

To summarize as simply as possible, evidence to date does not suffice to resolve the role of linkage relative to the observed ratios between estimates of dominance and additive genetic variance in hybrid populations. More evidence is needed. Direct comparisons of $F_2$ with $F_8$ or later generations in amounts about 3 times those already made are planned and should go far toward a conclusive answer.
E. Evidence concerning epistasis (interaction in effects of non-allelic genes) in corn populations.

Clear perspective on this issue demands careful recognition of the fact that a population may be completely homozygous at some loci and segregating at others. Completely homozygous means here that with reference to the locus in question there is only one allele present anywhere in the population. Epistasis may exist

1. Between genes of homozygous loci.
2. Between genes of one or more homozygous loci and those of one or more segregating loci.
3. Between the genes of two or more segregating loci.

Evidence outlined below pertains only to the third sort of epistasis. However, the effects of a larger group of loci are probably under scrutiny in studies involving material from the cross of two varieties than in studies confined to a single variety.

1. The regression of average phenotype on heterozygosity

Beginning with a pair of populations it is possible to prepare subsets of material which, aside from deviation due to sampling or selection, present an easily specified pattern of heterozygosity differences but are identical in gene frequencies. Then assuming

(a) no epistasis, and
(b) that non-genetic effects are random

it is not difficult to show that the expected relation between mean phenotype and average heterozygosity is linear. It follows that observation of significant deviation from linearity in experiments where absence of selection and random non-genetic effects are legitimate assumptions constitutes evidence for epistasis.
A group of experiments employing materials \((F_1, F_2, \text{backcrosses of three kinds, and parent lines themselves})\) derived from pairs of long-inbred lines was conducted during the period, 1948-1952. The range of heterozygosity (on a relative basis but also in absolute terms in the case of the homozygous parent lines) was 0, 25, 50, 75, 100. First backcrosses and the \(F_2\) provide two sorts of material at the 50 level. The results of these experiments are reported by Sentz \textit{et al}\ (Agron. J. \textit{46}, 514-520, 1954) and by Sentz (Ph.D. Thesis, N. C. State College); only a brief summary will be offered here. The regression of mean yield on heterozygosity was extremely variable from one set of material to another (the different sets derived from different pairs of inbred lines) and one trial to another with the same set of material. In some instances, the regression appeared to be essentially linear but in a number of cases the relative performance of 25\% heterozygosity material was high and of 75\% heterozygosity was low so that performance rose sharply from 0 to 25\% heterozygosity, was quite similar for the 25, 50 and 75\% levels and rose sharply again from the 75 to the 100\% level. Examination of the data on tasseling date revealed considerable differences among heterozygosity levels and a tendency for these to be associated with yield differences in a manner that might explain the non-linearity of yield (when it occurred). These differences in tasseling date served to indicate differences in development rate among the heterozygosity levels which in turn meant that environmental variation could not be assumed random. Thus the time pattern of weather conditions could well have been the source of variation in the pattern of results from one trial to another. With this in mind, we have never felt that the data could be interpreted safely with reference to the importance of epistasis. Perhaps the most
important conclusion was that data from such experiments can never be considered as very good evidence concerning epistasis unless there is valid reason to assume uniformity in the time pattern of development from one level of heterozygosity to another. While this is not a positive result, in view of the frequency with which the general sort of experiment is employed in studies of gene action, it is a significant one.

A second set of trials were conducted comparing two open-pollinated varieties, their $F_1$ cross, the $F_2$ from random matings of $F_1$ plants, and the backcrosses to both parents (bulked seed from crosses of random $F_1$ plants to random variety plants).

The work involved three sets of material based on the three possible pairs of three varieties (Jarvis, Weekley and Indian Chief). It consisted of two trials, one at each of two locations in 1955. This work was stimulated by results with D. pseudoobscura (Verthkov, Evolution 6:241-251, 1954). He found that when geographical races were crossed the $F_1$ exhibited heterosis, but the mean $F_2$ performance was more often than not below the average for the two parent races. The trait studied was larval survival. A logical interpretation (and the one offered by Verthkov) was that performance in the parents was the result, in considerable degree, of favorable epistatic gene combinations that were broken up by the recombination associated with production of the $F_2$ from the $F_1$. Mean yields in our trials with corn are listed in Table 4.18. Results for the two trials were very similar and hence only the averages for the two trials are presented. In each case heterosis was exhibited by the $F_1$; in all instances the $F_2$ and backcross means were intermediate between the $F_1$ and parent variety means; in no instance was there a significant difference between the $F_2$ and mean of the
Table 4.18. Mean yields for parent varieties, F₁, F₂ and backcrosses in three sets of material (averages for two locations).

<table>
<thead>
<tr>
<th></th>
<th>Jarvis, Weekley</th>
<th>Jarvis, Indian Chief</th>
<th>Indian Chief, Weekley</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of Varieties</td>
<td>.37</td>
<td>.36</td>
<td>.42</td>
<td>.383</td>
</tr>
<tr>
<td>F₂</td>
<td>.40</td>
<td>.39</td>
<td>.47</td>
<td>.420</td>
</tr>
<tr>
<td>Mean of backcrosses</td>
<td>.42</td>
<td>.38</td>
<td>.44</td>
<td>.413</td>
</tr>
<tr>
<td>F₁</td>
<td>.44</td>
<td>.40</td>
<td>.50</td>
<td>.447</td>
</tr>
</tbody>
</table>

bankcrossoes; and in no instance did the F₂ deviate significantly from the midpoint between the F₁ and the mean of the parent varieties. Unfortunately data were not collected on tasseling or maturity date to provide evidence on comparative development rates. Thus, repeatability of the results may be questioned on the grounds that what was observed may have resulted from a unique combination of differential development rates and the time pattern of environment in these two trials. In other words, the data do not establish that epistasis is unimportant in the differences among these varieties. On the other hand, they give no hint that epistasis is a significant feature of the variety differences.

2. Comparison (in random breeding populations) of the genotypic variance (a) among half-sib families, \( \sigma^2_m \); (b) among full-sib families within half-sib families, \( \sigma^2_f \); and (c) among individuals within half-sib families, \( \sigma^2_w \).

In populations that have been random breeding for many generations, genotype frequencies will be very close to those expected with independent segregation, i.e., the relative frequencies of genotypes will exhibit little effect of linkage. Then, if there is no epistasis
\[ \sigma_m^2 = \frac{1}{4} \sigma_g^2 \]
\[ \sigma_f^2 = \frac{1}{4} \sigma_g^2 + \frac{1}{4} \sigma_d^2 \]
\[ \sigma_w^2 = \frac{1}{2} \sigma_g^2 + \frac{3}{4} \sigma_d^2 \]
and \[ \sigma_w^2 - 3 \sigma_f^2 + \sigma_m^2 = 0 \]

On the other hand, the presence of epistasis would cause \( \sigma_w^2 - 3 \sigma_f^2 + \sigma_m^2 \) to be greater than zero, so that a check on the presence of significant amounts of epistasis is afforded.

In most of our D-I trials only \( \sigma_m^2 \) and \( \sigma_f^2 \) have been estimated but in three trials with open-pollinated varieties, an estimate of \( \sigma_w^2 \) was also obtained. This was done by including plots (randomly located) of single crosses (16 or 32) between inbred lines derived from the variety involved in the trial. Inbred lines from the same variety were used so that the single-crosses would constitute a sample of genotypes from the variety in question and would therefore tend to exhibit the same amount of non-genetic variance. The lines employed were as nearly unselected as possible so that the mean performance for single-crosses would be similar to that of the variety. This is important because of the possible effect of level of performance on variance. The difference between intra-plot variances of the varietal and single-cross material was used as the estimate of \( \sigma_w^2 \). Estimates of the three variances are listed in Table 4.17. These estimates are all predicated in definition of genotypic value with reference to the population of environments occurring in the year and location of the trial, i.e., they are biased estimates of \( \sigma_m^2, \sigma_f^2 \) and \( \sigma_w^2 \) for genotypic values defined relative to some broader population of environments (those
of a geographical area and a sequence of years).

If there is no epistasis, \( \sigma^2_I = \frac{1}{4} \sigma^2_g + \frac{1}{8} \sigma^2_d \). Then, assuming no epistasis, the total genetic variance is estimated as \( 4\hat{\sigma}^2_I = .00288 \) for the average of the three trials. Compared to this the estimate (.00028) of the average for \( \sigma^2_W = 3\sigma^2_I + \sigma^2_m \) is of trivial magnitude. On the other hand, the fact that the estimate is not one that merits great confidence is attested by the variation among the three estimates averaged.

Table 4.17. Estimates of \( \sigma^2_m \), \( \sigma^2_I \) and \( \sigma^2_W \) from three trials with open-pollinated varieties and single-crosses of inbred lines from the same variety.

<table>
<thead>
<tr>
<th>Variety</th>
<th>Year</th>
<th>( \hat{\sigma}^2_m )</th>
<th>( \hat{\sigma}^2_I )</th>
<th>( \hat{\sigma}^2_W )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis</td>
<td>1954</td>
<td>.00096</td>
<td>.00084</td>
<td>-.0013</td>
</tr>
<tr>
<td>Jarvis</td>
<td>1955</td>
<td>.00054</td>
<td>.00062</td>
<td>.0067</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>1956</td>
<td>.00020</td>
<td>.00069</td>
<td>.0011</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>.00072</td>
<td>.00072</td>
<td>.00217</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variety Mean</th>
<th>Mean of Single-Crosses ( \hat{\sigma}^2_W - 3\hat{\sigma}^2_I + \hat{\sigma}^2_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis</td>
<td>.27</td>
</tr>
<tr>
<td>Jarvis</td>
<td>.36</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>.52</td>
</tr>
<tr>
<td>Mean</td>
<td>.383</td>
</tr>
</tbody>
</table>
F. Responses to selection.

Five selection experiments have been in progress long enough to provide some evidence. In all five, selection has been among full-sib families and based entirely on grain yield. In four, selection has been based on average performance in two field trials (two replications at one location in each of two years); in the other, on performance in one trial (two replications). The field trials for family comparison were all in the Design I form so that each trial provided estimates not only of mean performance, but also of additive genetic and dominance variance. In all trials after those providing the basis for the first round of selection, check materials were included to provide bases for (1) estimating effect of selection on mean performance and (2) adjustments for inbreeding effects. The five genetic populations include two hybrid populations (21 x 7 and 34 x 45) and three varieties (Jarvis, Weekley and Indian Chief).

Evidence on the performance of material obtained by inter-breeding selected families is available for one cycle of selection in the three varieties and for two in the hybrid populations. The trial or trials on which selection was based provided the variance estimates used in predicting the effect on mean yield of the particular round of selection. Because of year to year variation in conditions affecting the level of yield, all predictions were put in terms of percentage of the mean yield. Table 4.19 compares observed and predicted increments after adjustment for inbreeding effects.

The evidence concerning response in the 34 x 45 population is ambiguous. Little progress from selection was predicted for either cycle; there appeared to be considerable improvement after one cycle, but after two the selected population yielded less than the F2 from which the population was initiated. Remembering that only rough accuracy of prediction can be anticipated in individual instances, the general
agreement between observation and prediction is very good indeed for the other four populations.

Table 4.19. Observed and predicted yield increments (percent of mean) in selection experiments.

<table>
<thead>
<tr>
<th>Population</th>
<th>Cycle of Selection</th>
<th>Predicted Increment</th>
<th>Observed Total Increment</th>
</tr>
</thead>
<tbody>
<tr>
<td>34 x 45</td>
<td>First</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>.3</td>
<td>4.3</td>
</tr>
<tr>
<td>21 x 7</td>
<td>First</td>
<td>9.7</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>5.1</td>
<td>14.3</td>
</tr>
<tr>
<td>Jarvis</td>
<td>First</td>
<td>15.5</td>
<td>15.5</td>
</tr>
<tr>
<td>Weekley</td>
<td>First</td>
<td>10.9</td>
<td>10.9</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>First</td>
<td>8.5</td>
<td>8.5</td>
</tr>
</tbody>
</table>

The effect of selection on additive genetic variances is also of interest. Present evidence is summarized in Table 4.20. The estimates listed are biased (by genotype-environment interaction variance) but since relative magnitudes are the prime issue and since, for each population, the bias is of constant kind, the estimates are pertinent to the issue in question.

Table 4.20. Estimates of additive genetic variance after different amounts of selection.

<table>
<thead>
<tr>
<th>Prior Selection</th>
<th>Population</th>
<th>34 x 45</th>
<th>21 x 7</th>
<th>Jarvis</th>
<th>Weekley</th>
<th>Indian Chief</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td></td>
<td>.0031</td>
<td>.0034</td>
<td>.0040</td>
<td>.0035</td>
<td>.0032</td>
<td>.0034</td>
</tr>
<tr>
<td>One cycle</td>
<td></td>
<td>.0028</td>
<td>.0020</td>
<td>.0008</td>
<td>.0045</td>
<td>.0008</td>
<td>.0022</td>
</tr>
<tr>
<td>Two cycles</td>
<td></td>
<td>.0013</td>
<td>.0021</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three cycles</td>
<td></td>
<td>.0002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The information available strongly suggests a moderate decrease in additive genetic variance as a consequence of selection. However, the known sampling error of the estimates and the variation in results from one population to another suggests caution in drawing conclusions.

G. Heterosis and genetic variation in variety crosses.

Materials obtained by crossing varieties are of interest as such but also for what we can learn from them about the nature of genetic differences among established varieties.

1. Heterosis in variety crosses.

Six North Carolina open-pollinated varieties were crossed in the 15 possible combinations. The six varieties and 15 F₁ crosses were compared in five replications at each of three locations in 1952 and 1953. The difference in mean yield between the F₁ and the average for its parent varieties was in all 15 instances in favor of the F₁. This observed heterosis ranged from 4.6 to 46.2 percent and averaged 19.9 percent for the 15 crosses. The Jarvis-Indian Chief cross is of special interest because of the extensive information being collected on this pair of varieties. The difference in yield between this cross and the average of the two varieties was 29 percent. Later estimates for this particular cross have been lower. When all are averaged, weighting by number of separate trials (in a different year, location or both), the average observed heterosis for this cross is 19.8 percent.

The one clear fact established by these observations is that the genetic differences among open-pollinated varieties are considerable; heterosis is not possible in the cross of two populations that are genetically the same. The nature of these differences or the reasons
for their being are not clear but must be adequately comprehended by a tenable hypothesis concerning the genetics of yield in corn.

2. The genetic variance among families in the $F_1$ cross of two open-pollinated varieties.

The genetic variance among half-sib families in the $F_1$ of a variety cross is one of the quantities required for predicting progress from recurrent reciprocal selection using the two varieties as the base populations in which selection for combining ability is practiced. To obtain estimates of this variance, Design I trials were conducted in which the experimental material derived from mating individuals of two varieties (Jarvis and Indian Chief). One set of families (256) were obtained by mating each of 64 Jarvis males (pollen parents) to 4 Indian Chief females; another set was made in the reciprocal cross mating each of 64 Indian Chief males to 4 Jarvis females. Each set of families was grown in two trials (two locations in 1954) and the entire work was repeated using new samples of families in 1955. The data provided estimates, for the $F_1$, of the genetic variance among half-sib families, $\sigma_m^2$, and that among full sib families within half-sib families, $\sigma_f^2$. These are presented along with the comparable estimates from earlier intra-variety studies with Jarvis and Indian Chief. In order to have all estimates on the same basis with respect to bias from interaction with year and location all are reported on an intra-trial basis, i.e., all reflect the total of variance due to interaction of genetic differences among families with year and location.

The most striking feature of Table 4.21 is the average difference between varieties and crosses in estimates of $\sigma_m^2$. There was a similar
Table 4.21. Estimates of $\sigma_m^2$ and $\sigma_f^2$ for Jarvis, Indian Chief and their reciprocal $F_1$ crosses.*

<table>
<thead>
<tr>
<th>Population</th>
<th>Year</th>
<th>$\sigma_m^2$</th>
<th>$\sigma_f^2$</th>
<th>$\sigma_f^2 - \sigma_m^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis</td>
<td>1950</td>
<td>.00098</td>
<td>.00150</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1951</td>
<td>.00146</td>
<td>.00195</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1952</td>
<td>.00067</td>
<td>.00040</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1953</td>
<td>.00133</td>
<td>.00137</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>.00111</td>
<td>.00130</td>
<td>.00019</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>1952**</td>
<td>.00077</td>
<td>.00071</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1953</td>
<td>.00083</td>
<td>.00173</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>.00080</td>
<td>.00122</td>
<td>.00042</td>
</tr>
<tr>
<td>Means for the varieties</td>
<td></td>
<td>.00096</td>
<td>.00126</td>
<td>.00031</td>
</tr>
<tr>
<td>Jarvis x Indian Chief</td>
<td>1954***</td>
<td>.00035</td>
<td>.00173</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1954</td>
<td>.00044</td>
<td>.00108</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1955</td>
<td>.00032</td>
<td>.00082</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1955</td>
<td>.00061</td>
<td>.00037</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>.00044</td>
<td>.00100</td>
<td>.00056</td>
</tr>
<tr>
<td>Indian Chief x Jarvis</td>
<td>1954</td>
<td>.00043</td>
<td>.00079</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1954</td>
<td>.00055</td>
<td>.00081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1955</td>
<td>.00049</td>
<td>.00012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1955</td>
<td>.00071</td>
<td>.00091</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>.00054</td>
<td>.00066</td>
<td>.00032</td>
</tr>
<tr>
<td>Means for the crosses</td>
<td></td>
<td>.00049</td>
<td>.00083</td>
<td>.00034</td>
</tr>
</tbody>
</table>

* Estimates listed for Jarvis and Indian Chief were those obtained prior to any selection since unselected varietal materials were used in the cross.

** Trials with Indian Chief were twice ordinary size so the total information was equal for the two varieties.

*** Trials in the same year were at different locations.
difference in mean estimates of $\sigma^2_f$ but since $\sigma^2_f$ can be shown to include a quantity equal to $\sigma^2_m$ and plus another associated with non-additive genetic effects, this difference is merely a reflection of that noted in estimates of $\sigma^2_m$. The average excess of $\sigma^2_f$ over $\sigma^2_m$ is very similar for the varieties and the crosses.

3. The genetic variances in the $F_2$ of a variety cross.

The genetic variation in the $F_2$ or later generation of a cross between two populations is a standard source of information on genetic differences between populations. Random $F_2$ plants from the Jarvis x Indian Chief cross were used as parents of families compared in a Design I experiment in 1956. Genetic variance estimates were $\sigma^2_m = .00210$ and $\sigma^2_f = .00129$. In view of the fact that the true value of $\sigma^2_m$ cannot be greater than $\sigma^2_f$, the relative sizes of $\sigma^2_m$ and $\sigma^2_f$ suggest that .0021 is an over-estimate of $\sigma^2_m$ in this population. Even so, the fact that the above estimate is 50 percent greater than the largest of six comparable estimates of $\sigma^2_m$ within Jarvis and Indian Chief themselves is worth noting. It appears that there may be considerably more additive genetic variance in the population derived from the variety cross than in either of the parent varieties.

H. Covariances among characters.

Character covariances have significance in terms of (1) selection and (2) the genetics of quantitative characters.

Both the genetic and non-genetic covariances between characters are pertinent to the choice of an optimum selection criterion (index). The object of selection in most corn improvement programs is simply increased grain yield. When this is so the issue of an optimum selection criterion boils down to whether information on traits other than yield can be employed to advantage in seeking to improve yield.
Genetic covariances reflect pleiotropy, linkage of genes affecting different traits or both. Thus the magnitudes and variation of genetic covariances in different populations and changes in magnitude with opportunity for gene recombination in new hybrid populations are potential sources of information on pleiotropy and linkage. Evidence for linkage also has overtones relative to numbers of genes affecting the character.

Analyses of covariance have been computed for most of the data that has been collected. However, to date the covariance component estimates derived have not been studied in a thorough way. Some attention has been given to the selection index problem. This has led to the conclusion that when economic worth is equated to grain yield, there is little to be gained and some risk of loss from attempts to use characters (other than yield) that we have measured in the selection criterion. It should be noted that we have not collected data on insect or disease resistance and the foregoing is not stated with reference to these traits.

Cursory examination of the evidence has revealed some population differences in genetic covariances and some changes associated with opportunity for recombination in hybrid populations. However, the total evidence has not been examined with sufficient care to allow evaluation or interpretation of those observations.
V. Experimental Results with Crops Other Than Corn.

A. Introduction

Most of this work has been conducted by workers outside the Experimental Statistics Department and with primary financial support by the N. C. Agricultural Experiment Station rather than the Rockefeller Foundation. There have been two exceptions. A Design II experiment with tobacco and a study for estimation of genotypic and genotype-environment interaction variances in Irish potatoes were conducted by members of this group. In other instances, we have participated in an advisory capacity; most of the work to which reference will be made was conducted in patterns that we suggested.

Crops with which some work has been done include tobacco, Irish potatoes, oats, peanuts, soybeans, lespedeza, cotton and strawberries. No effort will be made to report all the studies in detail. Special attention will be given only to aspects that have implications beyond the population and crop in question, i.e., that have a bearing on issues in quantitative and population genetics that transcend the single specie. These include (1) the general order of magnitude of genotypic variances in absolute quantity and relative to genotype-environment interaction variances (2) level of dominance in gene action in self-fertilizing species and (3) the magnitude of epistatic variance in self-fertilizing and asexually reproducing crops.

B. Magnitude of genotypic variance relative to advance from selection.

Characters studied included the primary product yield in all crops: seed yield of soybeans, oats, and peanuts; tuber yield in potatoes; berry yield in strawberries; forage yield in lespedeza and leaf yield in tobacco. Data have been taken on a variety of other traits, e.g., protein and oil content of seeds; pods per plant and amount of shattering in soybeans; number and size of berries in strawberries; lint length and strength in cotton; etc. The summary to follow will pertain primarily to yield characters.

Probably the most useful way in which to talk about the size of genotypic variance is in terms of the amount of improvement that seems possible through selection in the
light of the variance observed. Detailed interpretation would require specification of the exact bases for selection. Rather than go into such detail, an assumption is made that the comparison of selection units would be of accuracy that is roughly comparable or minimum, relative to ordinary procedures with the crop in question. For example, a randomized block comparison with two or three replications at from one to three locations in a single year.

The improvement expected as a percent of the mean was lower in the self-fertilized crops studied than in the two asexual ones, potatoes and strawberries. In the self-fertilized crops the estimates of yield improvement from selection among progenies obtained by selfing the F$_2$ plants from a cross of two pure lines were in the range from 0.5% to 15%. In strawberries and potatoes it appeared that the selection among seedlings from the cross of two clones should result in 30 to 40 percent improvement relative to the mean of all seedlings. However, the average for such a seedling population would probably be less (perhaps considerably less) than the yield of the parent clones.

In general the indications in all crops are that sufficient genetic variance exists to support moderate yield improvement in well designed breeding programs. Of course, what can be accomplished relative to yield is always moderated by the attention that must be given to other economic traits. The estimates of genotypic variance obtained were not in all instances free of bias from genotype-environment interaction variance. However, in a number of instances the data allowed separate estimation of genotypic variance and genotype-environment interaction variance and the figures that have been given for expected effectiveness of selection have been adjusted to take account of probable biases.

C. **Genotypic variance relative to genotype environment interaction variance.**

Information on this subject has been obtained with soybeans, cotton, lespedeza
and Irish potatoes. In the case of yield the estimate of total variance from interaction of genotype with year and location was considerable. It ranged from 60 to 160 percent of the associated estimate of genotypic variance. These figures are in line with the observations for yield in corn where it was estimated that the total variance from interaction of genotype with year and location is as large or a little larger than the genetic variance.

For certain other characters typified by plant height and oil content of seeds in soybeans, lint length and strength in cotton and tuber number in potatoes, the relative amounts of genotype-environment interaction variance were considerably lower, 40 percent or less of the genotypic variance. These compare in this respect with ear height and tasseling data in corn.

D. Level of dominance in gene action in self-fertilized crops.

Information has been obtained with one crop, tobacco. When estimates are obtained for populations derived without selection from the cross of two pure lines so that gene frequencies at segregating loci are close to one-half, the ratio of $\sigma_d^2$ to $\sigma_g^2$ can be interpreted relative to level of dominance. As noted earlier linkage may bias $\sigma_d^2/\sigma_g^2$ upward but low values are unambiguous in their implication. Design II studies were conducted with four hybrid populations of N. tabacum. No statistically significant differences between populations were observed in the estimates of either $\sigma_g^2$ or $\sigma_d^2$; hence, nothing will be lost by considering results in terms of pooled estimates for the four populations taken together. Data were taken on six characters: days to flower, plant height, leaf number, leaf length, leaf breadth and leaf yield. In the six cases, $\sigma_g^2$ ranged from 3.7 to 6.5 times as large as its standard error, i.e., there was without doubt a considerable amount of additive genetic variance in all traits studied. On the other hand, five of the six estimates of $\sigma_d^2$ were negative (not statistically significant) and the one for leaf length, while positive, was only
one-fourth as large as the estimate of $\sigma_e^2$ for that character (and not statistically significant).

The high proportion of negative estimates of $\sigma_d^2$ is a little disturbing. There is every reason to expect slight partial dominance which would mean a small positive value of $\sigma_d^2$. In any event, $\sigma_d^2$ cannot be negative. Given small real values of $\sigma_d^2$ one expects more positive than negative estimates rather than the reverse. However, there are no obvious sources of bias that would not also operate in the work with corn where a degree of dominance has been indicated for all characters studied. By comparison the evidence indicates less dominance in tobacco than corn and that, in fact, the level of dominance in gene action is very low in tobacco.

E. Epistasis as a source of genetic variance.

In self-fertilizing crops it is relatively easy to estimate the magnitude of genetic variance among families produced by selfing. An important issue is the extent to which this variance is due to additive genetic effects as opposed to non-additive effects. If any very large portion is non-additive, there are valid reasons for believing it due primarily to epistasis rather than dominance.

An experiment designed to provide information on the nature of genotypic variance among $F_3$ families in soybeans is now in progress. The approach being used is comparison of offspring on mid-parent regressions for (1) offspring produced by exclusive self-fertilization and (2) offspring produced by cross-pollination between two $F_3$ families followed by one generation of selfing for seed increase. The two kinds of offspring were produced using a common set of $F_3$ families as the parental material. The regressions are expected to be equal if all of the $F_3$ family genetic variance is from additive effects. If a portion is due to epistatic effects (or dominance) the regression is expected to be lower for the second class of offspring because the gene combinations responsible for epistatic effects are less likely to be transmitted.
intact in cross-fertilization. Data are available for two populations from trials conducted at two locations in 1956. The yield data have been analyzed. The smaller of the estimated parent-offspring regressions was for the first kind of offspring, those produced by only self-fertilization. This was unexpected since on any genetic hypothesis the regression for the first kind of offspring should be as large or larger than the other. The observed result provides no evidence for any non-additive effects.

In strawberries, asexual propagation allows exact replication of genotypes which greatly facilitates accurate estimation of the total genetic variance among individual genotypes. A Design II experiment with this crop was elaborated to allow estimation of the variance among individual genotypes within full-sib families \( \sigma^2_w \) as well as the genetic variance among half-sib families \( \sigma^2_m \) and that among full-sib families within half-sib families \( \sigma^2_f \). The joint interpretation of these three variances has already been discussed (Part IV, D, 2). In the strawberry data \( \sigma^2_w = 3\sigma^2_f + \sigma^2_m \) was large relative to the total genotypic variance. The result implies an important amount of epistasis though excess heterozygosis of the parental stocks employed may provide a part of the explanation of the observations. Characters to which these findings apply were number and average weight of berries and total yield of berries.

The final bit of evidence comes from early results in a selection program with cotton. Families produced by selfing after a cross of pure lines were grown together in a series of field trials. The data provided estimates of genotypic variance among the families and other variance components required for prediction of improvement to be expected from selection among the families. The families with highest lint yield were then inter-crossed and the population so obtained was field tested with a series of check varieties grown also in the original trials. The apparent improvement resulting from selection was about 70 percent of what had been predicted assuming all of the genetic variance among families to result from additive effects. The evidence indicates that most of the variance among families produced by selfing is additive.
VI. Listing of publications

A. Project Publications.


B. Publications on cooperative work.

Cotton


Lespedeza


Oats


Soybeans


Tobacco


Animals