ANALYSING AND PREDICTING SELECTION RESPONSE IN *TRIBOLIUM*

By

William Bunball

A THESIS

Submitted in partial fulfilment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Sheep Husbandry Department
Massey University

November 1966

I N D E X

TABLES

FIGURES

INTRODUCTION

This thesis is concerned with comparing experimentally, some of the methods of analysing and predicting the responses of populations to artificial selection.

The thesis involves two experiments, of which the first one aims at supplying data to compare in detail two systems of measuring the response with regard to its average rate, its fluctuations and its trends, and to compare several systems of predicting this response. The results of this survey are then tested to see how well they apply in the second experiment, which is run far more along the lines of a commercial breeding project.

There are many reasons for undertaking this subject of research but perhaps the main one is the author's opinion that existing literature contains fairly few examples of selection experiments whose results have been subjected to, or indeed are amenable to, a detailed analysis. Without such analysis the results are far less informative than they could be. But more important than this is the fact that an analysis of the selection response is indeed necessary to throw light on the genetic processes causing such response, and the better the analysis the greater will be the understanding of these genetic processes. Conversely, detailed analysis will serve to test how well current genetic theory does manage to explain or predict selection responses. From the viewpoint of commercial plant and animal breeders it is important that the selection theory on which their projects are based, be readily available, valid, and easily applied. Years of work and thousands of pounds could be wasted by choosing the wrong selection procedure in a project or indeed the wrong project altogether. Certainly there is a huge amount of literature available on selection theory but the more important point is - how well is it able to explain the results of an actual project and how well does it predict them? Such a test of selection

theory has been made often before now, with varying results. The aim of this present thesis is virtually to discuss how valid these past tests themselves have been, and to conduct a more searching test of its own. It will not be implied that the results necessarily extrapolate beyond the present data.

This aim is executed in the following sequence of work.

PART ONE: ANALYSING SELECTION RESPONSE

The section begins by surveying the various methods of analysis used in the past, as reported in the literature. During this literature review there is special emphasis laid on two aspects of analysis; firstly the extent to which each earlier experiment was *equipped* to provide data for a full analysis, and secondly the extent to which this data was actually *exploited* in the analysis. This inevitably involves a discussion of the items considered to be essential in the carrying out of any selection experiment, and of the procedures considered to be advisable in the analysis of results. The final part of this section then describes the present author's selection experiment, which deliberately contained these essential items and was analysed as suggested. It involves selection for pupal weight on two populations of *Tribolium castaneum* Herbst. Each population is selected for both lighter and heavier pupae, and the experiment lasts for eighteen generations.

PART TWO: PREDICTING SELECTION RESPONSE

Once again the investigation begins with a survey of the literature, specifically dealing with the prediction of response to selection pressure. Discussion of the various methods and especially of the reliability of their stated success, gives no absolute guide to the author on the validity of prediction theory: he therefore retests the standard methods of prediction (and some other

ones besides) on the *Tribolium* populations already under selection.

PART THREE: EFFECTS OF SELECTION INTENSITY

This section aims to test the above findings under new conditions of selection procedure, conditions chosen for similarity to those of commercial breeding projects. Selection is carried out in one direction only (for heavier pupae) but from a broader base population and at three separate intensities. There is a more refined mating system used for the selection lines but proportionately less effort is spent on the Control population. The results of this experiment are then analysed and also assessed for predictability, by the methods used in the earlier experiment.

It is stressed that these three sections of the thesis form a naturally-continuous project, and it is only for simpler presentation that they are written up as separate units. Each section is just one aspect of the larger aim - which is to assess the usefulness of mathematically-derived selection theory from the viewpoint of its application to commercial breeding operations, and if possible to improve this usefulness. Since there is still a large gap between the two aspects despite the attentions of persons with far more time and facilities than those of the present author, it might be argued that this thesis has little chance of closing such gap. However, the validity of selection theory will be proven not by one large and outstandingly successful experiment, but rather by the accumulation of results from many experiments on diverse organisms and under diverse conditions. The one criterion that all these experiments must have in common is that each of them be on its own a truly valid and reliable test of the theory. It is this criterion which the author will endeavour to show has been missing from several earlier experiments, and which he endeavours to rectify in the present thesis.

ESSAY ON SELECTION THEORY

This thesis is concerned with the analysis and prediction of selection response, especially in measuring their accord with selection theory. Obviously one of the main pre-requisites for such a study is a detailed review of selection theory itself and the following essay aims to provide such information. The historical development of this theory will be traced only briefly but its present state will be described in detail. However, there is a huge amount of literature available on selection theory and so only the aspects relevant to this thesis will be emphasised.

CONCEPT OF ADDITIVE VARIANCE: The concept of additive genetic variation is the central aspect of quantitative selection theory. A long time ago, but unnoticed for many years, Mendel (1865) had shown experimentally with garden pea that some gene alleles are dominant to others at the same locus. Consider the diagram below, which represents a population varying at only one locus (which has two alleles, A and a).

genotype value:	+a	d	-a
genotype frequency: (at equilibrium)	p^2	2pq	q^2

The value (d) of the heterozygote depends on the degree of dominance, increasing from zero where there is no dominance to a value greater than "a" where there is overdominance. The population mean value is calculated from the value x frequency of all genotypes and if selection operates so as to

increase the frequency of any genotype, then the population mean will change by a calculable amount. But the above system does not adapt very well to selection theory because those genotypes chosen to be parents are in fact broken down during gamete formation, and their genes are inherited in new combinations. What is needed instead is a system that will measure the values of genes rather than those of genotypes. Such a measure exists and is called the "average effect of a gene"; it is calculated as the average deviation from the population mean of individuals which received that gene from one parent, the gene received from the other parent having come at random from the population. In the above case the average effects of gene alleles "A" and "a" are respectively $q\left[a+d(q-p)\right]$ and $-p\left[a+d(q-p)\right]$. The sum of these average effects over both alleles in an individual is called its "breeding value" (or "additive value") which when applied to a parent expresses the mean value of the progeny. The difference between this breeding value in an individual and the genotypic value given earlier is called "dominance deviation" and it obviously represents the non-heritable portion of genotypic value.

The existence of more than two alleles at a locus complicates the mathematics of this partitioning approach but introduces no new principles. If however there is more than one locus varying in the population then the genotypic value of any individual includes not only the additive values and dominance deviations summed over all loci but also now an interaction between them. This non-allelic interaction is termed "epistasis" and like the dominance deviation it is a property of genotypes rather than of genes, so that its value is liable to alter at each generation. Thus three separate portions - additive, dominance and epistatic - make up the genotypic value of an individual. Unfortunately major environmental factors such as nutrient availability and minor ones such as error of measurement may mask this value when measured from the phenotype.

To summarise -

phenotypic value = additive value + dominance deviation + epistatic deviation + environmental deviation.

When this partitioning approach is applied to all individuals of a population then the statistical variance of phenotypic values can be represented as

phenotypic variance = additive variance + dominance variance + epistatic variance + environmental variance.

There may also be a covariance between the environmental component and any of the other three but those other three genetic components are by definition uncorrelated with one another. In addition there may also be a component of interaction variance, caused by interaction between environment and any genetic component.

Of all four components only the additive value depicts the heritable portion of an individual's phenotypic value (in Lush's narrow sense of heritability) and likewise only the additive variance depicts the heritable portion of a population's phenotypic variance. The ratio of this additive to phenotypic variance is therefore called the "heritability" of the selection character, though its value of course applies strictly only to that particular population in that particular environment at that particular time. If the heritability of a given selection character is high (which means that environmental and other non-additive effects are small) then the phenotypic value of an individual is on the average a fairly good indicator of its true breeding value.

EFFECT OF SELECTION: The basic effect of selection is to increase the frequency of certain genes and this implies that there will be an increase of homozygosity. It also means that the additive variance of the population will be lowered, and with it the heritability and future selection response. As selection continues the population will move toward homozygosity for these favoured genes (or genetic

drift will fix their undesirable alleles) and both the heritability and response will tend to zero, becoming asymptotically zero when the population is fully homozygous at all of those loci affecting the selection character. Any variability then remaining in the population is due to the differential effects of environment on individuals, and these effects are not heritable.

The description above would be called an additive interpretation of selection response, for it assumes that the effect of a gene is not greatly altered by dominance and epistatic effects of other genes. But where over-dominance exists in a population the heterozygous individual will be favoured in selection ahead of either homozygote, and as the frequency of heterozygotes is duly increased the selection response will again fall off. Gradually an equilibrium point is reached at which the loss of homozygotes by non-selection each generation is equalled by their gain from random recombination of gametes from the heterozygous parents. At this equilibrium point, which Falconer (1961) says is determined more by the relative merits of the homozygotes to one another than by relative superiority of the heterozygote, genetic variability is still present in the population but there is no response to further selection. Epistasis has much the same effect as overdominance; that is, any individuals whose phenotypic superiority depends on epistatic gene combinations will have these combinations broken down at gamete formation. The response to further selection will gradually taper off, becoming zero when the superior epistatic combinations break-down as fast as they are selected. If at any time the selection pressure is removed then this continuing break-down of combinations will cause the population mean value to regress, perhaps eventually back to its original value. In the overdominance case such regression would probably be less severe, though reversed selection would eventually become very effective. In the additive case neither relaxed nor reversed selection would affect the

mean of a homozygous limit population.

CALCULATING THE GENETIC COMPONENTS OF VARIATION: The discussion above shows that it is theoretically possible to explain the genetic nature of a given limit population in terms of additive, dominance, and epistatic gene effects. Likewise it is theoretically possible to predict the nature of such a limit population if the relative importance of these three effects is known beforehand. This is only one of many good reasons for trying to determine the additive, dominance and epistatic variances of a population even before selection starts; another good reason concerns selection procedure - if the additive variance was estimated to be very low it would predict a poor response from using many standard methods of selection. However, it is usually impossible to estimate exactly these three components because such estimates depend on knowing the value and frequency of each gene involved, and also the epistatic relations among them. Most genes involved in quantitative selection characters have effects far too small to be isolated in this way and instead therefore the three components are calculated indirectly, from the degree of resemblance observed between related individuals in a population. Except when caused by non-genetic effects common to members of a family, and different between families, such as cage effects or maternalism, a high degree of resemblance between relatives indicates high heritability and hence also high additive variance. By this method the observed components of phenotypic variance - within and between families - can be equated to the causal genetic components given above. For example a halfsib family contains by definition a mean genotypic value equal to half the breeding value of the common parent. Their covariance is the variance of the means of all halfsib families and therefore equals the variance of half the breeding values of all parents; this in turn equals a quarter of the additive variance in the population.

Thus Cov (HS) = $V_A/4$

This estimate of additive variance is actually biased upwards by one sixteenth of the additive by additive epistatic variance and by successively smaller amounts of higher-order epistatic variance, but the amount of bias will usually prove to be very small.

Different methods of partitioning the observed phenotypic variance produce different kinds and amounts of bias. This point is discussed in detail by Dickerson (1959), though he assumes throughout that there is no artificial or natural selection operating and that therefore the mating of parents and survival of their progeny is entirely random. Dickerson explains that the least biased estimate of heritability is obtained from regression of offspring mean on midparent. The estimate can be written as

$$h^2_{\overline{op}} = b_{\overline{op}} = (V_A + \tfrac{1}{2}V_{AA} + \tfrac{1}{4}V_{AAA} \text{ etc} + \tfrac{1}{2}V_M + 1\tfrac{1}{2}Cov_{MG} + V_L)/V_P$$

where V_{AA} = additive by additive epistatic variance

V_{AAA} = additive by additive by additive epistatic variance

V_M = variance from genetic variation in direct maternal effects

Cov_{MG} = covariance between total genetic deviation in the transmitted and the direct maternal effects

V_L = genetic variance of sex linkage

V_P = phenotypic variance

The estimate is subject to bias also from genotype-environment interaction and environmental correlation between parent and offspring, but such effects are usually difficult to measure. Positive assortative mating of the parents will greatly increase the midparent variance and therefore reduce the sampling error of estimate, but unfortunately this could raise a consider-

able bias if individual gene effects are not small and if the population contains much non-additive variance (Reeve, 1961). Dickerson states also that the above regression estimate is valid strictly only in populations where each litter has a different sire and dam. When each sire has several mates, as is often the case in animal breeding projects, heritability should be estimated instead from the regression of offspring mean on sire and dam separately. The estimates could then be combined providing correction is made for any phenotypic correlation between mates. The estimate of heritability from offspring-dam regression exceeds that from offspring-midparent regression by an amount of $(\frac{1}{2}V_M + Cov_{MG})/V_P$, though of course this bias could be negative if Cov_{MG} is large and negative. Offspring-sire regression is rarely used by animal breeders because commonly each sire is mated to more than one dam and so the estimate from regression on sire has a larger sampling error than that on dam.

The main alternative system of estimating heritability uses mating layouts similar to those for regression analysis, but involves a variance analysis of the progeny alone. This progeny phenotypic variance can usually be subdivided into three components; between sires, between dams within sires, and within dams (fullsibs). Of these three the sire estimate contains the least bias; it may be written as

$$h^2_{4S} = (V_A + \tfrac{1}{4}V_{AA} + \tfrac{1}{16}V_{AAA} \text{ etc} + 2V_{LQ} + V_{GE})/V_P$$

where V_{GE} = variance due to genotype-environment interaction.

This method deviates from the regression estimate of offspring on midparent by an amount of

$$(V_{GE} - \tfrac{1}{4}V_{AA} - \tfrac{2}{16}V_{AAA} \text{ etc} - \tfrac{1}{2}V_M - 1\tfrac{1}{2}Cov_{MG})/V_P$$

It is potentially quite a large deviation but the actual size and direction of this bias depend mainly on the genotype-environment interaction and the maternal

effect. On the other hand heritability estimation from the dam component involves the extra correlation among fullsibs, so that it is further biased upwards by dominance variance, fractions of epistatic variance containing dominance components, and non-genetic effects common to fullsib litters. But the third possibility, estimation of heritability from phenotypic correlation within fullsib groups alone, involves by far the greatest bias. This estimate exceeds that from offspring-midparent regression by an amount of

$$(\tfrac{1}{2}V_D + \tfrac{1}{4}V_{AD} + \tfrac{1}{8}V_{DD} \text{ etc} + \tfrac{1}{2}Cov_{MG} + 1\tfrac{1}{2}V_M + 2V_C + \tfrac{1}{2}V_{L\sigma^7} + \tfrac{1}{6}V_{L\varphi} + V_{GE}) / V_P$$

where V_D = dominance variance

V_{AD} = additive by dominance epistatic variance

V_C = variance due to common environment for fullsib litters.

In all regression and variance methods used above the standard error of heritability estimate depends on the number of sires, dams, or progeny measured. Standard errors are an important item to consider when choosing between the several estimation methods given above because for example the decision whether to select an individual on the basis of its own value or of its family mean depends as much on knowing precisely the upper limit of heritability estimate as in having an unbiased but imprecise estimate. This question of estimate precision becomes more important as fewer individuals are to be measured, for it then becomes vital to make the best use of these individuals. The question is discussed fairly fully by Robertson (1959) but since it does not bear directly on the thesis it will not be repeated here, except for two comparisons. For the case when heritability equals 20%, halfsib variance analysis has a precision about equal to that of regression analysis of offspring on one parent. Below 20% sib analysis is generally preferable, above 20% regression analysis is better; assuming an equal total number of individuals to be measured in

each case. Secondly, regression on midparent is apparently always more precise than regression on either one parent.

It is clear from the equations given above that epistatic variance can be subdivided into units of as high an order as desired, but there are two reasons for not subdividing too far. Firstly, selection techniques are seldom advanced enough to make use of such refinement, and secondly each new variance component resulting from subdivision requires one more partitioning equation to evaluate it. If estimates are desired for each of four component variances - say additive, dominance, epistatic and environmental - then the mating system must be organised to give four appropriate partitioning equations of the types above. Simultaneous solution of these four equations will then give the four estimates required. It is better of course if there are more equations than unknown components, for then it is possible to measure the errors of estimate of these components: in such cases the equations are solved by methods of least squares (Mather, 1949b) or maximum likelihood (Nelder, 1960; Hayman, 1960). Mather's method is straightforward but unfortunately takes no regard of the differences in precision with which the separate covariances are estimated. Nelder and Hayman showed that this weakness could be mostly removed by weighting the estimates according to their standard errors; but this requires iterative calculations, which may be too laborious without a computer. Cooke et al., (1962) compared the two approaches using hybrid Drosophila stocks and reported little difference between them, the unweighted estimates of variance components being 60-98% as efficient as the others in terms of the size of sampling error. Weighting is of course unnecessary if the variance estimates are roughly known beforehand, for in that case the degrees of freedom can be adjusted to equalise the observed phenotypic variances of the various progeny groupings; but it is not usual to possess this prior information.

PLANT vs ANIMAL APPLICATION: The partitioning theory described above applies equally to plant and animal populations but the estimation of additive variance etc using this theory is carried out slightly differently in the two cases. Animal breeders usually set up a hierarchial mating system in which each sire has a different harem, and the resulting halfsib and fullsib progeny are analysed as shown above. The book called "Animal Breeding Plans" by Lush (1943) is perhaps the standard work in this field. The modifications which plant breeders apply are mostly improvements, for example much higher replication of progenies. Furthermore, pasture plant breeders are usually able to take an almost limitless number of tillers from any one plant and thereby measure environmental effects precisely using many identical genotypes. Animal breeders cannot usually afford high replication of progenies except with the small laboratory species, and secondly replicated genotypes such as identical twins are not easily identified. The factor of low cost also enables plant trials to be laid down in several areas at once thereby giving an assessment of the genotype-environment interaction: this is a variance component now considered to be of major importance in any estimation of heritability (Hanson, 1963). On the whole plant breeders have been able to devise the more complex and informative experiments for predicting heritability, for example Designs I, II and III of Comstock and Robinson (1948) and in particular the diallel cross (Sprague and Tatum, 1942). This latter design involves all possible matings among a set of parents chosen randomly from the population. The resulting progeny (including the reciprocal crosses and often the parents as well) are then grown and measured in randomized replicated blocks, preferably over more than one year and location. Such a design would involve several years of matings if applied to animals which have a long generation interval and only a few progeny each cycle. It is an interesting point that the use of a diallel cross to evaluate

General Combining Ability (GCA) and Specific Combining Ability (SCA) of various plant genotypes was common long before Sprague and Tatum interpreted them in terms of variance components. Gardner (1963) tabulates the variance components and their mean square expectations for all designs above but stresses the assumptions inherent in them - that there are no non-genetic maternal effects nor multiple alleles present and that there is linkage equilibrium, no selection pressure, and regular diploid behaviour at meiosis. He also discusses the results obtained by applying these designs to breeding projects with corn, alfalfa and forage grasses. In many cases the plant material is highly selected and so the results cannot be extrapolated to a wider population. In other cases the experiments are poorly planned for estimating the genetic parameters with equal precision. Kempthorne and Curnow (1961) discuss some of the problems involved here; for example they show for a partial diallel cross that the degrees of freedom can be adjusted between GCA mean square and SCA mean square to make either one more precise - but the ratio of these estimates and hence of the heritability itself may suffer in the process.

Self-fertilised plants can also be assessed for heritability by a diallel cross. Matzinger (1963) records that this usually involves the various progeny generations resulting from a cross between two pure lines and it has to be assumed that there is no epistasis present and that the alleles at each locus have a constant frequency of 0.5 through all generations.

Finally, many commercial plant species are polyploid. In alfalfa, apparently autotetraploid, several characters are still inherited in a diploid manner but this property is often hard to determine for any given character. Allopolyploids of course usually behave as diploids with regard to chromosome segregation. However, the statistical genetic theory applicable to true polyploid behaviour is very complex, which explains why it has received little

attention in the literature. For example Comstock and Robinson (1952) describe circumstances in which even inter-allelic interaction contributes greatly to the genotypic variance among halfsib families. Kempthorne (1957) also discusses how polyploidy may bias different covariance estimates to different degrees, but he does not probe the subject very deeply.

HISTORY OF SELECTION THEORY: The essay so far has discussed the present state of partitioning theory based mainly on additive, dominance, and epistatic variances. The historical development of this theory is not so important to the thesis but it will be considered briefly.

The three components of variance were first defined in 1918 by Fisher, who then went on to partition the covariances of simple relatives in terms of them. Wright (1920) also studied the subject experimentally with guinea-pigs but Cockerham (1954) extended the mathematical approach to cater for non-random mating systems and multilocus epistasis. More important, he subdivided the two-locus epistatic variance into four fractions - additive by additive, additive by dominance, dominance by additive, and dominance by dominance. By a slightly different approach Kempthorne (1954) was able to generalise further by catering for an arbitrary number of alleles at each locus, and more recently (1963) to include even several non-random mating systems in this unified approach. Occasionally new mathematical treatments and terminologies have been offered (Malecot, 1948; Mayman, 1955) but they have not generally proved popular.

The particular importance of variance partitioning to selection theory was also realised a long time ago. It began with a short and simple treatment by Wright (1921b) on the change of mean value caused by selection in a population of known genotypic array, and this approach was expanded in a series of special case papers by Haldane (1924-7). The third important work about

this time was Fisher's "Genetical Theory of Natural Selection" (1929) - important because his Fundamental Theorem of Natural Selection would adapt so well to the case of artificial selection. This theorem states that "the rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time" (page 37). Fisher denotes fitness by the Malthusian parameter (m) which is based on the relative growth in numbers of different genotypes in the population. His theorem is developed in terms of known gene frequencies and values and as with most genetic theory at this early stage it was soon realised to be oversimplified. Fisher himself observed (1941) that the model was valid only for those mating systems in which the genotype frequency ratio of (heterozygote)^2 / (homozygote AA x homozygote aa) stays constant, such as in random mating. Crow (1955), Wright (1955), and Crow & Kimura (1955) all tried to increase the generality of this theorem, to cater for changes of natural fitness in time, for variable selection coefficients, for random sampling variations, and for small mutational and migrational effects. Kempthorne (1957) also recognized the value of this theorem but he could not follow Fisher's condensed derivation by Malthusian parameters. He therefore re-derived from first principles an analogous equation catering for <u>artificial</u> selection but this derivation is rather tedious because it involves matrices. A simpler derivation was given by Kimura in 1958. It was based on a quantity called "coefficient of departure from random mating" (Θ) and could cater for an arbitrary number of loci and alleles, also for non-random mating systems. Kimura's equation for measuring the change of yield under artificial mass selection is

$$\Delta\bar{Y} = \sum_{ijkl\ldots}(\Delta Y_{ijkl\ldots})P_{ijkl\ldots} + 2\left(\sum_i \alpha_i \Delta x_i + \sum_k \beta_k \Delta y_k + \cdots\right) + \overline{\Sigma\epsilon . \log \Theta}$$

where the first RHS term is the effect of an improving environment,

second RHS term is the additive variance for yield,

third RHS term is the joint effect of gene interactions and changes in the mating system. This term involves a dominance-inbreeding factor and an epistasis-linkage factor.

OTHER FACTORS AFFECTING SELECTION RESPONSE: Kimura's equation above is perhaps the present peak of all attempts to cater for the many parameters affecting selection response, and it illustrates very well that such response involves more than just the additive, dominance and epistatic variances. Among the extraneous factors consider firstly the effect of linkage.

Because it involves complex mathematics linkage has been studied little until recently. Geiringer (1944) proposed a system of parameters that would completely describe the linkage relations among an arbitrary set of loci but the system was far too unwieldy to have much applied use. Cockerham (1956) showed for a population in coupling-repulsion linkage equilibrium that linkage affects only the epistatic variance (increasing it by an amount that depends on the recombination index) and therefore biases the covariances of fullsibs, halfsibs, and cousins, but not those of parent-offspring. But Cockerham could not devise a simple method of evaluating any particular bias. Such a method was suggested recently by Schnell (1963) and although it becomes complex to use when more than three loci are treated together, he argues that the partitioning of variance is seldom carried to such lengths in practice anyway.

Schnell's paper dealt with populations in linkage equilibrium but most theoretical and experimental studies on linkage are concerned with disequilibrium populations (Mather, 1949b; Comstock and Robinson, 1952; Gates et al., 1960) such as the progeny of a cross between two homozygous parents. Such crossing programmes are common in plant breeding work and in them the progeny population is expected to be in very severe linkage disequilibrium -

and the breeder will probably not afford to wait the several generations of random mating required to restore the equilibrium. Jones (1960) has recently formulated in general terms the linkage effects following such a cross of two homozygous parents. The flexibility of this approach is considered by Hayman to be a breakthrough on the theoretical front but it is not yet in a form useful to experimentalists.

The extra complications that arise when selection is superimposed on linkage in populations should also be mentioned. Lewontin and Kojima (1960) showed for a two-locus model that if epistasis is present then the genetic response is affected by the linkage value. If this linkage is tight enough there may be permanent linkage disequilibrium. Similar results are arising with more complex gene models (Lewontin, 1964) and it now seems clear that linkage can affect both the limit of response and the rate of reaching it. Unfortunately there is apparently no general recipe for measuring this effect and it appears that linkage at present can be handled confidently only when a computer is available.

Secondly consider population size as a source of bias to selection response. A small population could introduce bias in two ways, one directional (inbreeding depression) and the other non-directional (genetic drift). If only a few parents are selected each generation then inbreeding depression can be expected to occur with any characters related to fitness - and this means almost all characters important in modern commercial breeding. The more closely related to fitness the greater will probably be its depression per unit of inbreeding and this rate relationship will be linear if no epistasis is present and selection is not operating. Robertson (1955b) explains how conversely the degree of dominance and of inbreeding depression are usually very slight in characters such as Drosophila bristle number, which character is unlikely to be important to the fitness and survival of

the genes.

Even if the selection character does not display inbreeding depression the use of too few parents increases the genetic drift that may occur (though of course need not occur in any given case), according to the formula of Crow (1955). He calculates that for a population of randomly mating monoecious diploids, the variance of drift in gene frequency ($V_{\delta q}$) equals

$$\frac{q(1-q)}{4N} \cdot \left[1 - F' + (1+F')\frac{V_k}{\mu_k}\right]$$

where q = frequency of gene

N = total number of offspring

μ_k and V_k are the mean and variance of the number of surviving progeny per parent

F' = inbreeding coefficient of parent generation.

Therefore as the population number falls the gene frequency fluctuates more greatly and with it, the population mean value and response to selection. A different approach to the same source of bias is pointed out by Robertson (1961), in which the selection process itself may reduce the effective population number. If the heritability of selection character is fairly high then the individuals chosen to be parents will be more closely related (that is, come from the same family) than will random members of the population; hence the effective population size will be less than that calculated from the actual number of parents.

Finally of course, if only a small population is measured each generation then sampling error may become important, sufficiently so to bias the estimate of selection response.

The third factor which may bias a selection response is scale, and this time the direction of bias can be determined. Tests for any of the many different types of scale may be carried out on response data but the one type

most likely to influence the selection character of body weight in this thesis is that scale effect caused by the exponential nature of growth. This effect results in a correlation between the population mean and its variance; such that the variance ought to increase under selection for increased mean weight (and this is usually borne out in practice) - and hence also the selection differential and selection response should increase also. This effect may be reduced partly by the fact that such selection would simultaneously decrease the amount of genetic variance remaining but nevertheless it is logical to suspect a scale effect on such occasions as when two-directional selection produces an asymmetrical response, greater in the upward direction than downward. The test usually applied for this growth scale effect in cases of asymmetry is to convert the data to logarithms. An approximate conversion formula based on the population mean is given by Wright (1952)

$$\overline{(\log X)} = \log \bar{X} - \tfrac{1}{2} \log (1+C^2)$$

where X = arithmetic value

C = coefficient of variation.

Any scale effect which is due to the exponential nature of growth will disappear by converting the data to logarithms. Note however, that the converse is not necessarily true; that is, the disappearance of asymmetry under such treatment need not imply that the asymmetry was caused by scale - it could have instead been due to directional gene dominance, natural selection, or inbreeding depression (Falconer, 1961).

Fourthly consider the effect of natural selection on a population response. This force operates against the response in the sense that the more extreme individuals in a set of parents produce fewer mature progeny than do the less extreme ones. It can operate in two ways - either natural selection is against the extreme phenotypic values themselves, or it is against the more

homozygous individuals which under the additive concept of selection contain these extreme phenotypes. These two aspects have been discussed by Robertson (1956) and Latter (1960). In each case it was shown that the bias caused to selection response depended on the strength of natural selection for intermediate genotypes or intermediate phenotypes.

PREDICTING SELECTION RESPONSE: The essay so far has been concerned with describing in turn all those forces which might affect a selection response. If such response is to be predicted beforehand (and in most commercial stations it is surely important to make these predictions before choosing between potential projects) then the effects of these forces must also be predicted beforehand. This is usually impossible to do more than qualitatively, for the reason that the selection process itself alters these forces as it proceeds. Even the additive variance changes, and its rate of decline can be predicted only when the frequencies and values of all genes are known (Mei, 1963). Strictly speaking therefore, a prediction of response is valid for one generation only, but fortunately it usually seems to hold good for quite a few generations (this longer time is apparently needed before the additive variance has declined, or forces such as natural selection have built up, to a degree that seriously alters the rate of response). Nevertheless the methods of estimating heritability given earlier - namely regression analysis and sib analysis - are still open to serious bias over even one generation, for they assume that natural selection, scale, and inbreeding depression etc are not operating. This bias is not effective if the prediction applies to selection response averaged over opposite directions, but most commercial projects involve unidirectional selection only. In these cases an asymmetrical response is not catered for. The one prediction system, apart from an actual generation of previous selection, which does claim to cater for asymmetry does not involve variance estimates

at all. It is put forward by Abplanalp (1961) and is called a "linear heritability estimate". The system involves selecting prospective parents from a pedigreed population and determining their average genotypic and phenotypic values. These genotypic values are calculated by least squares analysis of displacement effects (displacement of a sire mean from all-sires mean, of a dam mean from all-dams-within-sire mean, of an individual value from all-progeny-within-dam mean) instead of from variance estimates. Then the genotypic values of prospective parents are compared by regression with those of the remaining population. Abplanalp explains how this system is superior. It does not assume (as variance analysis does) that the regression of genotype on phenotype is a straight line and secondly it relates much more closely to actual selection procedure. It is clearly a system that justifies test in this thesis.

By whatever system the heritability is predicted in a population, its estimate must be multiplied by the value of selection differential to give the predicted response. Fortunately the differential can usually be measured directly from the population. It has already been explained that although such predicted response theoretically is valid for only one generation it is usually quite safe to extrapolate the actual observed response for a few generations further ahead. However, some attempts have been made to refine the prediction of a response to long-term selection. In 1960 Griffing studied the case of truncation selection based on the individual phenotype. The response after n generations of selection was calculated to be (for two loci)

$$\mu_n = \left(\frac{i}{\sigma_p^2}\right)\left\{n\sigma_A^2 + \sum_{r=1}^{n}\left(1-y\right)^{r-1}\cdot \tfrac{1}{2}\sigma_{AA}^2\right\}$$

where i = selection differential in absolute units

σ_P^2 = phenotypic variance

y = recombination index.

This equation predicts that the response will decrease in succeeding generations as the coefficient to epistatic variance falls, but this epistatic contribution will be lost if selection is relaxed at any time and the superior gene combinations are allowed to break up. Griffing's equation can cater for arbitrary dominance, unequal gene recombination in the two sexes, and any number of alleles; but it disregards natural selection, it disregards the change in additive variance itself, and it disregards the genetic drift which would occur in finite populations. Gill (1965b) points out that the response is thereby greatly overestimated, this discrepancy increasing as the prediction becomes longer-range.

Another type of long-term prediction equation, though less sophisticated, is given by Robertson (1960). He calculated for the case of individual selection that the expected limit of response equals about $2\sigma_A^2 N_e \bar{\imath} / \sigma_P$ in other words about 2 N_e times the response at the first generation (where N_e = effective population number). This equation caters only for additive genes and thus assumes that the response limit is reached only when additive variance is exhausted. Compared with Griffing's equation it ignores epistasis and linkage. Robertson admits these deficiencies of his model but implies that its value lies in the ease of handling it mathematically. He suggests that such parameters as linkage and epistasis can be handled with confidence only when a computer is available.

ANALYSING SELECTION RESPONSE: The essay earlier described in turn all those forces which might affect a selection response. But from a practical view-

point exactly the opposite approach is needed; that is, given a particular response curve, how does the breeder determine which of these several forces have influenced the shape of the curve? This analysis should be attempted for three reasons. Firstly, if there had been attempts to predict the response and they had failed it would be valuable for future reference to know why they failed. Secondly a knowledge of all the forces operating would guide the breeder in his selection system from then on. Thirdly this analysis may provide basic knowledge on gene action which, though not of immediate use to the breeder, may be of value to the geneticist.

The first requirement in analysing a selection response is undoubtedly to separate the genetic and environmental causes of change in population mean. This calls for a reliable Control in the project. From one point of view the Control should be unchanging in genotypic array from generation to generation for then any phenotypic fluctuations it may show must be due to changes in the environment. Such a genetically constant Control requires that there be no pressure from natural selection, that mating procedure be kept strictly random to prevent any unconscious artificial selection, and that the population be kept large enough to avoid bias from genetic drift or sampling error. From another point of view however, such a Control of constant genotype is imperfect in that it must inevitably diverge from the line under selection. Once these two populations begin so to differ in genotype there is no guarantee that they will be affected similarly by changes in the environment (Bray et al., 1962). The obvious way to avoid such bias from genotype-environment interaction would be to take a random sample from the selection line each generation and allow it to merely case-mate for one generation. But a Control of this nature still confounds the effects of environmental change with those of natural selection. The only practicable way to meet this dilemma is apparently to extract a

subline each generation as above but then to select it in the opposite direction instead of simply mass-mating it. The divergence of means between this Control and the main selection line measures fairly accurately the genotypic response combined of both lines but it loses information on the bias caused by natural selection, scale, and environmental fluctuations, because it cancels these effects instead of measuring them. To sum up, it would seem possible to measure all these effects separately only if more than one type of Control could be used but this is a luxury not usually afforded in selection experiments. Occasionally commercial projects do include a complex Control. For example Goodwin, Dickerson and Lamoreux (1960) outline a repeat mating scheme for poultry which measures environmental changes from inter-year comparisons of different progenies from the same parents, measures genetic changes from intra-year comparisons of progeny groups from two or more successive generations, and even measures changes in maternal effects associated with ageing of the dam. This scheme is biased only by sampling error of the various progenies.

Despite their weaknesses listed above, Controls are usually credited with removing much of the environmental fluctuation from a phenotypic response, though sometimes this credit is based on preconceived ideas as to what shape the genetic response should be. This is particularly so when the response surface (after adjustment for the Control) traces an asymptotic curve, for such a curve is to be expected under the additive, dominance and epistatic models of gene action, also when natural selection begins to exert increasing pressure on the changing mean, and finally also simply under the effect of a deteriorating environment.

If such an asymptotic curve is obtained it is possible to analyse for some of the extraneous effects, at least qualitatively. For example if the

mean does not regress when selection is relaxed then natural selection against it is likely to be absent. If the population does not respond to reversed selection it may be homozygous, which indicates the absence of either epistasis or overdominance. If however, the population does regress when selection is relaxed it probably indicates the break-up of epistatic combinations or the effects of natural selection against increased mean or increased homozygosity, or perhaps all three. The deconfounding of these three effects requires careful genetic analyses and perhaps extra Controls. The effects of epistasis on selection response are so similar to those of overdominance that there is still some controversy over which one is more important in natural populations (discussion in Falconer, 1961 : pages 287-291).

Analysis of a selection response should include attempts to evaluate the effects of sampling error and genetic drift. Both of these items depend on population size and there are formulae available to estimate their expected effects, but such formulae cannot of course show if these full effects are actually operating nor the direction of their bias. It is therefore difficult in a given case to prove that the trends or fluctuations of a response curve are due to drift and not to natural selection say, but this difficulty is mostly overcome by carrying a number of independent replicated selection lines. The random effects of drift should be cancelled out when considered over all replications. The more replications are present the more likely is this cancellation, but unfortunately high replication is not usually afforded in commercial breeding projects.

VALUE OF COMPUTERS: It should now be clear that the many different forces affecting a selection response are very difficult to handle for either predicting or analysing such response. On the one hand there is an experimental problem of being able to afford enough replications, Controls etc to even make the

separation of these forces potentially possible and on the other hand there is an analytical problem of being able to carry out the long and complex processing of data that would usually be necessary. But for both prediction and analysis the solution to these problems is being approached in an indirect way - by the use of computers.

Computers can carry out the random Monte Carlo processes that are basic to genetic theory and they can handle complex gene models quickly. For simulating a selection project that is either envisaged or is already completed, a computer programme is written containing parameters for each property of the population concerned such as size, additive variance, genotype array, and linkage relationships. Those parameter values which are known and those unknown ones which must be guessed are included in the programme, along with instructions that subject the "population" to "selection" for any number of "generations" and with any number of "replications". Then the values of unknown parameters are varied, in the case of prediction until a broad range of such values is catered for, in the case of analysis until the simulated and actual response curves are sufficiently alike. The genetic models themselves in the simulated and actual populations are then accepted as being alike, though this evidence is always circumstantial of course.

The value of computers for simulating a response or for predicting a range of probable responses, still depends greatly on the detail and accuracy with which the parameters of the actual population itself are estimated beforehand. If the computer's gene model includes a wrong estimate of additive variance in the actual population then the prediction will surely fail. It is likely also that genetic models written for computers to date are greatly oversimplified. But as this skill at devising models increases, and also the skill of measuring parameters in actual populations, it is easy to foretell an

incredible use of computers for analysing and predicting selection response curves. Already past experiments have been re-analysed. For example Mather obtained in 1943 a temporary response plateau in his selection experiment with Drosophila. His explanation, based to only a small extent on genetic analysis, was that the resumption of response after plateauing was due to a change in the linkage relationships of the genes concerned. Fraser (1962) simulated the same experiment on a computer and he reports that Mather's explanation is plausible, for controlled linkage adjustments did indeed produce this same effect. Lewontin has similarly used computers to test the interpretation of data from earlier experiments (Lewontin and Dunn, 1960; Lewontin and White, 1960; Lewontin, 1962). On the prediction side computers have shown to be more useful in comparing the relative expected responses of two alternative selection procedures than in determining the absolute response from any one procedure. For example Martin and Cockerham (1960) compared by computer analysis the responses to selection at two intensities. They showed that for the particular gene models tested, more intense selection gave a measurably greater response in the early generations but was eventually overtaken by the response from weaker selection. The degree of superiority between intensity treatments varied as the gene model was altered.

Computers were not used in the present thesis because of the restriction of time available.

To conclude this essay it should be noted that no mention has been made of topics such as family selection or selection for an intermediate optimum value. There are two reasons for this; firstly the essay is orientated towards the actual thesis subject and secondly such topics, even if they were relevant

to the thesis, do not involve any big new principles of genetic theory. Those extra minor points of genetic theory that do become relevant will be discussed in the appropriate context later.

PART ONE: ANALYSING SELECTION RESPONSE

A main aim of this thesis is to study what use of selection theory can be made, and has been made, for analysing a selection response. The previous essay gives some idea of the first part, namely the amount and type of genetic theory that could help in interpreting such a response. But the full use of this theory requires elaborate experimental setups containing high replication, detailed and frequent genetic variance analyses, reliable Controls, and perhaps other subsidiary non-selection lines. These are luxuries that most commercial breeders cannot afford, and understandably so. It therefore rests upon laboratory breeders to provide such detail in their projects, for it is only from actual populations and experiments that genetic theory can be validly tested. An adequate experiment is therefore the first requirement of such a test; the second requirement, and it is equally important, is that the resulting data be analysed fully and validly. There is no doubt that the innumerable selection experiments reported over the years are considered to give general support to the current selection theory. It is the intention of this next part of the thesis to take a brief look at some of these experiments themselves, to see how well they fill the two validity requirements given above. There is no attempt made to deliberately pick holes in these experiments nor to discuss only those of them which are clearly unreliable; but instead merely to point out the range from good experiments to bad ones. Following this literature review will be reported the thesis experiment, which was designed to overcome some of the obvious weaknesses in past experiments.

1.1 LITERATURE REVIEW:

The two separate aspects of analysing a selection response should be clearly recognized. The first aspect is a practical one; it concerns the

incorporation of devices such as Control and replication into the actual project. The second aspect, and it clearly is limited by the first, concerns the analysis of data from the main selection line and from these various devices. Obviously the more simple is an experiment the more limited is its analysis. In its simplest form a selection project did not involve either replication or Control (Goodale, 1938) but historically it was quite early that Controls in the form of reversed selection (Chapman, 1946) or non-selection lines became standard procedure. Replication in the sense of different populations being subjected to simultaneous and identical selection pressure was illustrated by F.W. Robertson (1955) but true replication in the sense of completely independent lines drawn from one population is perhaps best shown by Clayton et al., (1957). Perhaps the final main step in sophistication to note is the extraction of non-selection and reversed selection lines at intervals during selection or the crossing of separate lines (Mather and Harrison, 1949; Robertson and Reeve, 1952).

The second aspect of a selection response - analysis of data - has also developed in sophistication. Perhaps the simplest form of analysis is to merely graph the successive generation means and interpret their pattern, but often regression lines or more complex curves are also fitted to such graphs. Also, where the experiment permits it, parameters such as the variance between replications can be analysed and are of considerable value. Mather and Harrison (1949) selected for a higher and lower number of abdominal chaetae in Drosophila populations. There was only one replication in each direction but subsidiary lines were extracted at various times. The authors explained much of the main line response in terms of these subsidiary line performances but it seems unfortunate that there was very little tabulation of response data. It is not denied that their long and careful discussions are very

valuable to the reader, for example in explaining the relation between response, variability, and linkage but it is suggested that this paper's conciseness could probably have been improved with the presentation of more data in tables. In many respects the paper of Reeve and Robertson (1953) illustrates this point. These authors likewise extracted subsidiary lines for relaxed and reversed selection during the experiment and they also added tests for inbreeding effects, for presence of lethal genes, and for parent-progeny relationships. Most data from such tests is summarised in tables, which not only makes for quicker reference but also gives the reader a better understanding of the authors' methods of analysis and reasoning.

Sometimes an author's methods of analysing data are perfectly clear but fairly simple. For example Kyle and Chapman (1953) merely record selection response as the total population response divided by the number of generations of selection. For their purposes this mean response over a period was sufficient description for they desired only to compare such observed result with a predicted response over the same period. But were such predictions to go astray it would be useful and much easier to learn why they did so, if there was more data presented on fluctuations and trends of response within the period. Falconer (1953) was able to use such interim data in his selection experiment for large and small size in mice. From the error variance of the mean response he subtracted the fraction common to both upward and downward lines, that is the covariance, resulting from environmental fluctuation. Next he subtracted the fraction due to sampling error. The remainder (which in his case was almost none) could then be attributed to genetic drift. This variance analysis by Falconer did not make use of data from all generations available but it is nevertheless very enlightening and indicates that considerable thought was given to the analysis of data by him. Another example of this thought is his

decision to measure the mean response by linear regression rather than from the usual formula, total response divided by number of generations. His reasoning and its validity will be discussed later but at least the sophistication was a big step forwards. The main weakness of Falconer's work was on the experimental side - it contained only one replication in each direction.

The analysis of selection response carried out by Thoday and Boam (1961) is fairly incomplete in some respects. Their experiment on Drosophila is subtitled "Description of Responses" but this description is presented as a long discussion unrelieved and unenlightened by any tables at all concerning the actual response to selection. The danger of this non-analytical approach is shown by one of their conclusions. They suggest that three particular lines "showed remarkably similar patterns of response" (page 175) but they support this claim only with a graph on which the three response surfaces are superimposed. Moreover these three lines do not have similar histories, that is they are not replications of one another. Nor do these similar patterns of response even refer to similar generations of selection.

Clayton et al's study (1957) of chaeta number in Drosophila was much more detailed. It involved several selection intensities and mating systems, and it was replicated up to five times. But Clayton too was mainly concerned to test prediction theory and so he apparently did not bother to analyse the observed response beyond this simple need; he merely averaged the responses from all generations and replications. There is no data given on the variation of response between replications even though it was admitted to be large. However, the authors did try to measure the various forces concealing the true genetic response. This they did by extracting non-selection subsidiary lines and by crossing plateaued lines, and most of these results are presented in tables of data.

One experiment which was both constructed and analysed in more detail than

most is reported by De Fries and Touchberry (1961). These authors carried out selection for Drosophila body weight at two intensities, in two directions, and with ten replications. Although most replications had unfortunately died out by Generation 10 the authors had collected sufficient data to calculate the confidence intervals of mean response and to give estimates of the inter-replication and intra-replication variance components. The analysis showed that there was great variation of response between replications. This was not surprising because each replication was a fullsib family and genetic drift would be near its maximum under such a mating system. However, the authors did report that a similar degree of variation under less severe inbreeding has been observed by Marien (1958), also with ten replications. The main value of the De Fries and Touchberry paper then is perhaps that they show very clearly how the analysis of response can be taken beyond the population parameter of mean response and into the population parameter of variance. Not only does this second parameter make it easier to interpret the first, for example in matters of scale; but it is also able to lead to information impossible from the first, for example the distorting effect which selection may give to the distribution curve of phenotypic values. If this curve becomes too far distorted from a normal distribution it becomes less meaningful to describe the phenotypic variability in terms of variance or standard error.

This short literature review shows a range of thoroughness in carrying out and analysing selection experiments. Many experiments are good in some respects but poor in others; for example an author may have ensured adequate replication and Control but given just a superficial analysis of the results. But perhaps the most striking weakness in previous experiments is that the analysis was usually concerned only with the mean response per generation; there was apparently very little interest in measuring the trends and fluctuations between generations which are hidden in this mean, nor of studying the

variability between replications. It is the present author's belief that much useful information is thereby lost which could have led to a clearer understanding of the forces involved in a selection response.

The first experiment described in this present thesis aims to correct these weaknesses. Its structural design is intended to allow a detailed analysis of response, and this detailed analysis is subsequently attempted. Before the experiment is described however, the methods and materials used in this thesis will be explained.

1.2 MATERIALS AND METHODS:

ORGANISM: Tribolium castaneum Herbst. (flour beetle) was the organism used in this thesis. It has nine pairs of autosomes and apparently there is equal crossing-over between them in male and female (Tribolium Information Bulletin, 1960). Tribolium is simple to raise and cheap to provide with good food and climate. In good conditions it is prolific (15 eggs per female per day) long-lived (one year) and has a life cycle of less than one month. It was preferred over the mouse because of this shorter life cycle. Recent extensive use of Tribolium by the Bell school at Purdue University discloses no serious faults for population selection studies. Its main drawbacks are apparently cannibalism and diseases but these can largely be controlled by management, and they were no problem in the present thesis.

The selection character used in this thesis is weight of the Tribolium pupa. It was chosen because it was known to respond well to selection and because it was easy to measure. It was also chosen because pupa weight is a "commercial" character in the sense that most commercial breeding projects are concerned at some stage in selecting for bigger plants or animals, because such bigger individuals yield more of the desired product. However, this does not imply that the results taken from Tribolium body weight can automatically be applied to the body weight selection programmes of plants and higher animals, and the dangers involved in such extrapolation will be discussed at the end of the thesis.

The Tribolium stocks used came from Ruakura Agricultural Research Centre at Hamilton; and originally from the Institute of Animal Genetics at Edinburgh University. The stocks had no special marker genes but their early selection history is not fully known.

Drosophila, so popular in the past for genetic research, has some dis-

advantages for this type of experiment. It contains only three pairs of autosomes and there is no crossing-over in the male, therefore linkage changes could have big and erratic effects as they apparently did in Mather's experiment reported earlier. These effects would unnecessarily complicate the analysis of response. Secondly, such a small chromosome number is probably never found in plant and animal species used for commercial breeding; in which case the results from Drosophila may extrapolate very poorly (Lerner, 1958).

DESCRIPTION OF TRIBOLIUM LABORATORY: This description will be given in detail because the lack of assistant labour or of expensive facilities required that the techniques of management and measurement had to be unusually efficient. This inevitably created a fair amount of unorthodoxy, which in turn is often considered a danger to validity.

Most of the apparatus was borrowed from the nearby Grasslands Division of D.S.I.R. and indeed the first few generations of selection were carried in incubators at that Division. Subsequently the laboratory was established on the Massey University campus and the D.S.I.R. apparatus was transferred there and re-calibrated.

It is not intended that the description given below should serve as a model for any general Tribolium research, for this was not in the author's mind at all when the laboratory was set up. It was instead set up to do just the one particular job - enable the selection programme of this thesis to be carried out.

(a) Apparatus: The experiment was housed in a Qualtex incubator (Model 324) which was water-jacketed, unlighted, and of internal size 30" by 24". It was fitted with a thermostat but while the experiment was running at D.S.I.R. this thermostat once broke down and allowed over-heating, thereby killing all

animals inside the incubator. (The experiment was restarted from reserve matings). No similar problems occurred when the experiment was shifted to a newer, though identical, incubator at Massey; however, it would be good policy to fit an over-riding thermostat before starting any future experiments. A separate incubator housed all the reserve matings of this thesis and although it was not an identical model it could be calibrated to produce the same climate inside.

The only other big item of apparatus was an old Baird and Tatlock weighing balance which with its moving vernier scale could weigh to 0.1 milligrams (pupae weights ranged between 0.6 and 4.4 mgms during the experiment) and this was nearly always precise enough to identify the heaviest or lightest pupae for selection. However the repeatability of this balance was poor at the lower end of its vernier scale especially on cold days, so only the middle range of the scale was used for weighing, and the laboratory was always warmed with a wall heater before weighing began for the day. But the main innovation was in the method of weighing. The conventional weighing method required the worker to switch on the electrical scale, wait for it to steady, place a pupa on the pan and wait for the new scale reading to steady, switch off, and then remove and store the pupa. The author could not afford time for this ritual nor could he afford the inevitable wear on this complex switch mechanism, for the total project required a weighing of more than 80,000 pupae. Therefore a system of "weighing by differences" was devised. Pupae were dropped one by one on the scale pan and their weights noted one by one from the successive changes in scale readings. A large cardboard map of the scale pan was placed alongside the balance and as each pupa was dropped on the pan and its weight noted, its position was marked on the map by a small card which also recorded that weight. When all 32 pupae had so been weighed the author merely had to scan the map, refer from it back to the pan, and so collect the heaviest (or

lightest) animals for selection. The 32 card weights were then just recorded in a book and the cards returned to their pool. The success of this system also required that the laboratory be kept very warm, otherwise the pupae tended to roll about on the scale pan and more time was spent chasing them than weighing them.

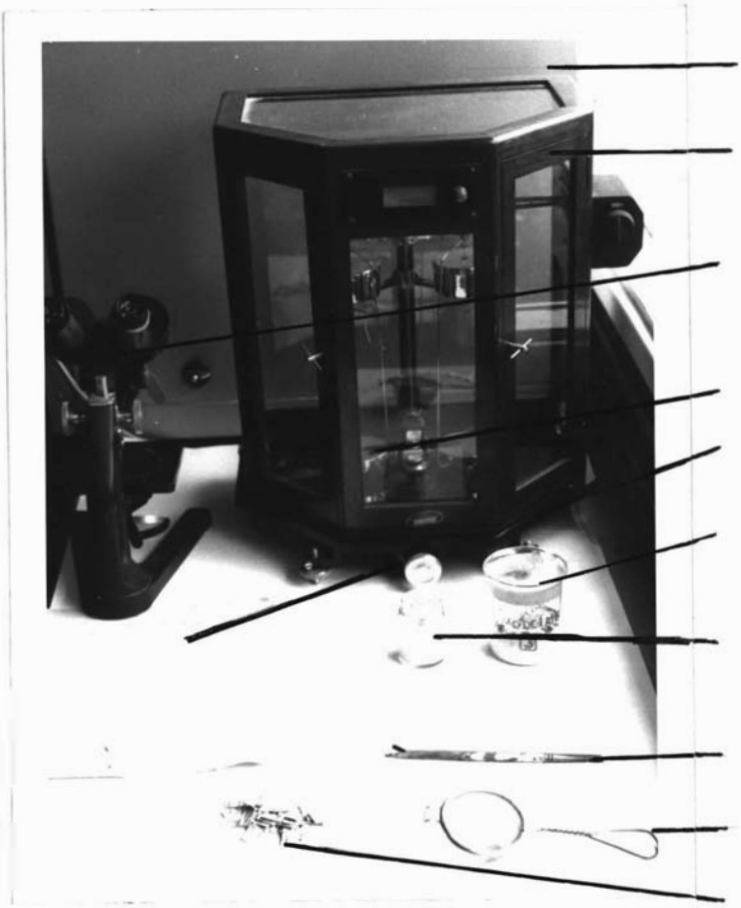
A binocular microscope was used for identifying the sex of pupae. The female has two prominent genital lobes on the ventral surface of its terminal posterior segment, just in front of the cerci. The male has not. When, as in the diallel cross the adults themselves had to be identified, they were sexed at the pupal stage then isolated until they matured into adults. Then the males were painted with a yellow spot on their backs. This system (Wong, 1964) proved by test to be more reliable and less unpleasant (to animals or author) than the older system of etherisation.

The eight-parent matings of Experiment One were kept in 250 ml beakers, a replication to each beaker. The pair-matings of Experiment Two were kept in 50 ml glass jars except for the Control pair-matings, which were kept in 10 ml vials. All of these containers were left open at the top to allow better circulation. Since the beetles were never observed to fly during the whole thesis this system was considered to be quite safe provided the container walls were kept sufficiently smooth to prevent them from crawling up.

Pupae were separated from their substrate using a small kitchen sieve. This did not enable eggs and small larvae to be separated out also, but it was not necessary to do this during the experiment anyway. The only other apparatus (Plate 1) was assorted glassware, coloured tape and chinagraph pencils (for identifying containers), a paintbrush (for transferring animals) and a killing jar.

When any replication was ready for weighing, all pupae in the container were sifted from the substrate and gathered in a petri dish. Then they were

Plate 1 : APPARATUS



randomly sampled and sexed until the quota of 32 females and 32 males was reached. Towards the end of the project some replications became too weak to provide the required 32 pairs at any one sampling. In such cases the replication was re-sampled about three days later; in the meantime all pupae already weighed were kept apart from one another in vials, and either selected or killed when the full quota was reached. Finally, pupae lose weight as they mature, therefore neither very young pupae (whiter than normal) nor very mature pupae (browner than normal) were sampled for weighing.

(b) Climate: Inside the incubators a climate is considered to be optimum when it allows Tribolium to have the shortest life cycle consistent with a high survival rate. This optimum is reported to be about 32° Centigrade and 70% Relative Humidity (Howe, 1956). The desired temperature was got by electrical heating of the incubator water-jacket and maintained with a thermostat. The humidity was got and maintained by placing a big tray of saturated salt (NaCl) solution on the incubator's bottom shelf (Solomon, 1951). This gave an R.H. value of about 77%. Checks with a hydrothermograph occasionally during the project showed that this climate varied by only a small amount, most greatly when the incubator door was opened. In such a climate the life cycle is approximately: egg for three days, larva for fifteen days, pupa for four days, and immature adult for four days. In a test described below, adults could produce enough eggs in this climate to provide several pupae per female per day.

(c) Diet: The standard diet for Tribolium in selection experiments is a ratio of sifted wholemeal flour to dried brewers yeast at 95:5. This is also close to the optimum diet for fast development and high survival (Tribolium Information Bulletin, 1960). However, commercial brewers yeast is expensive

and subject to import control so in order to get a substitute, the author obtained moist cakes of the yeast from a local brewery, then freeze-dried and crushed them. The success of this substitute was tested by a trial described below.

Preparation of the food is also fairly standard. The correct amounts of flour and yeast were weighed out then blended. They were ground through a sieve of 60 meshes per inch, sterilised at 70°C overnight, then resifted. This food was always prepared in small quantities, to keep it fresh.

A trial was set up to compare the food-values of three types of yeast - commercial brewers yeast, "live" DYC yeast, and the home-made brewers yeast described above. Each was made up into the above diet, eight beakers of food per yeast and in each of the beakers was placed two female and two male pupae. These animals mated after maturing and a few weeks later their progeny began to pupate. Thenceforth each day the newly hatched pupae were counted and removed. In Table 1 each value given is the sum of pupae for all eight beakers in that yeast treatment and no attempt is made to calculate the significance of differences between either beakers or treatments.

Table 1: Diet Trial (Units: new pupae per day)

Day after matings	22nd	23rd	24th	25th	26th	27th	28th	29th	30th	Total
Yeast component										
DYC "live"	25	28	63	64	92	138	138	66	39	653
commercial brewers	1	5	11	30	47	70	81	64	34	343
home-made brewers	43	57	82	92	78	134	112	84	55	737

The table shows clearly that the home-made yeast is apparently superior to both other yeasts, as well as being cheaper. More important, the superiority is most marked in the early stages, which means that a replication will produce its quota of pupae considerably sooner than in the other diets. The advantage of being able to so shorten the life cycle determined that this home-made yeast would be used throughout the present thesis.

There was evidence from the literature that cannibalism was a major danger in Tribolium and that it was greatly increased by a shortage of food. If all treatments were not given ample food then those that produced the heaviest pupae by selection would be penalised for they would have less food available per unit of bodyweight. Cannibalism would therefore be greater in such treatments and so a new factor of natural selection would be present, biasing the results. It was thus decided that all animal stocks be given food in excess of that needed by the biggest of them. That is, all replications would be given the same amount (an excess) regardless of whether they contained heavy stocks or light ones. But to determine just how much food was ample for the heaviest stocks was difficult and the literature (e.g. Rich, 1956) was very little help because it usually dealt with only one stage of the life cycle. The resolution finally adopted in this thesis after tests for evidence of cannibalism (such as half-eaten corpses) and tests for discolouration of the food was that the eight-parent matings of Experiment One be given 70 grams of food, and the pair-matings of Experiment Two be given 20 grams except for those of the Control, which were given less.

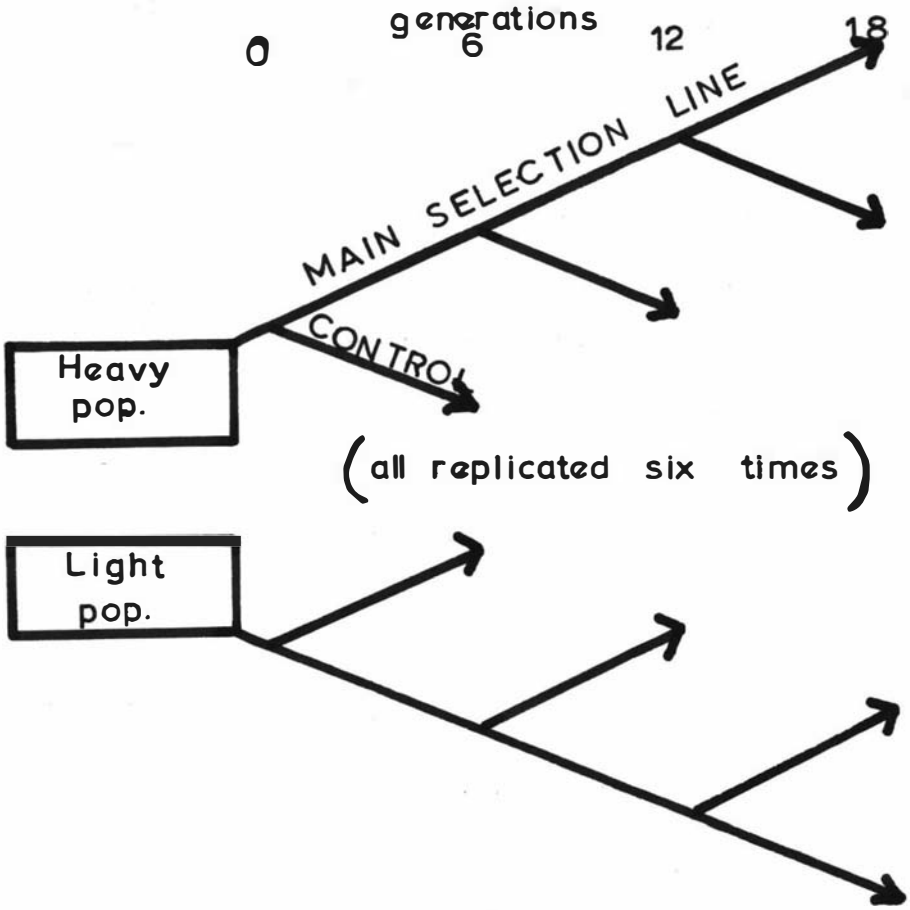
(d) Diseases: The prevention of disease was based on hot-water sterilisation of glassware and instruments after handling each replication. The laboratory floor and big items of apparatus were disinfected about monthly. Dummy jars of food were placed through the incubator and laboratory. The continued

apparent absence of diseases or stray beetles in them suggests that the precautions above were adequate. No mould was observed on the food at any stage of the project.

Lastly, the welfare of all cultures was checked twice weekly. The value of this practice was mainly to substitute the reserve matings for any cultures that had died. This occurred only with the pair-matings of Experiment Two, and only infrequently in these.

This then is a brief description of the Massey Tribolium laboratory. It has been inspected by Professor A.E. Bell who heads the Tribolium research programme at Purdue University, and was declared by him to be valid and adequate, though very unsophisticated.

PROGRAMME: Whilst the Tribolium stock was held at Manukura a sample of it was subjected to selection of moderate intensity for higher and lower pupa weight. Some progress was made in both directions over eight generations, resulting in "Heavy" and "Light" sub-populations. It is these Heavy and Light stocks that were used as base populations in the present thesis experiment, and they were used in the following way:



Selection of moderate intensity was attempted for heavier pupae in the Heavy population, and for lighter pupae in the Light population. It was continued for eighteen generations - the maximum time available during the thesis. This total time was subdivided into three periods of six generations each, namely generations 1 to 6, 7 to 12, and 13 to 18. This procedure enables the three periods to be compared and adds more interest to the analysis because the periods correspond roughly with what may be regarded as the early, middle, and late stages respectively of a long-term commercial breeding programme.

CONTROL: Another big advantage from subdividing the total eighteen generations available is that it allows new Control lines to be set up at the start of each shorter period. This reduces the chance for a selection line to depart far from the Control line in genetic content, and it therefore reduces the bias that may result from genotype-environment interaction (Bray et al., 1962).

Considerable care is needed in choosing the type of Control to be used. The earlier discussion on this subject showed that no one type of Control is apparently able to give a measure of each influence such as scale, natural selection, and environmental change. A Control which can evaluate one of these influences is usually unable to evaluate the rest. A decision must be made therefore as to whether the Control shall be used to help evaluate the true genetic response or instead to evaluate one of these subsidiary influences that bias measuremen of the genetic response. The former role was decided upon, and the best type of Control for such a role is actually a selection line in the opposite direction from that of the main line, and of equal intensity (Falconer, 1961; page 198). Hence, each replicate selection line in the present experiment has a mirror-image Control and the true genetic response of such a replication is best estimated by averaging the separate responses of selection line and Control. This is the case because any influences such as natural selection, scale or environmental change which bias the main selection line response in a certain direction will bias the Control response in the opposite direction and (approximately) by an equal amount. The effect of averaging is thus to cancel such bias, leaving an unbiased estimate of the true genetic response. This system of using a Control obviously loses information on the effects of these subsidiary influences but fortunately (and this influenced the decision) there are sometimes independent ways of evaluating them other than from the Control; for example the influence of scale can be studied also from

the changes of variance in the selected line itself.

REPLICATION: In any experiment the degree of replication which is optimum depends on the relative sizes of between-replication variance and within-replication variance. This is a ratio which is likely to alter as selection progresses but no information could be found in the literature which would give a guide on the correct ratio for Tribolium experiments. Choosing the degree of replication therefore became more a matter depending on the personal interest of the author in this aspect, though of course a sensible balance must be kept between the number of replications and the number of animals in each one. Before any further progress could be made on this subject, the total size of the experiment would have to be fixed and this in turn depended on two points - the work which could be handled by one person, and the distribution of this work between the two experiments of the thesis. The experience of workers with Tribolium at Ruakura was not a very strong guide because it soon became obvious that the weighing procedure could be streamlined and that the life cycle could be shortened. The amount of work eventually decided on for this Massey thesis was settled after a short trial. (It subsequently proved to be rather too much work at times).

There is another way also of determining how many animals should be weighed each generation in order to detect treatment differences. It is based on the formula from Cochran and Cox (1950, page 20):

$$r \geq 2\left(\frac{\sigma}{\delta}\right)^2 (t_1 + t_2)^2$$

where r = number of animals required per treatment

δ = true percentage difference that it is desired to detect

σ = true standard error per animal (expressed as percentage of the mean)

t_1 = significance value of t in the test for significant differences between treatments (2-tailed)

t_2 = value in t - table corresponding to 2(1-P),

where P is the desired probability of obtaining a significant difference between treatments.

The term "treatment" in this experiment refers to the mean response over all replications for any one generation in either direction and if provision is made to compare such treatments in pairs should the occasion arise, this provides 1 (r -1) degrees of freedom for the estimate of error. By assuming in advance that r would be quite large the degrees of freedom was set at ∞ for purposes of the t - table. Following this, t_1 and t_2 were set at 2.576 and 1.645 respectively, and δ was set at 5. The composite meaning of these three values is that a standard of precision has been set for detecting treatment differences and this standard is - that 5% differences between treatments will show out with 95% certainty on a 1% significance test. The final parameter in the above equation, σ , was determined from a pilot survey of the unselected Ruakura base population. In this survey (which had other uses as well) 576 random pupae of each sex were weighed and the results were as follows:

	FEMALE	MALE
mean weight (μ)	= 1.99 ~~ngms~~	1.89 ~~ngms~~
standard deviation (s)	= $\pm$ .236	$\pm$ .245
coefficient of variation (s/μ)	= 11.9%	12.9%
standard error of mean ($s/\sqrt{n}$)	= $\pm$ .0098	$\pm$ .0102

The coefficient of variation was taken to be 12%, which gives σ = 12 for the equation.

Substitution for all terms in the equation gave the solution r = 205.3.

However, a new problem had now arisen. The mean weights of female and

male pupae were compared using the t-test and resulted in the solution d = 7.1 (very highly significant). This means that the female pupae are distinctly heavier than the males and so must be weighed and analysed separately; or at least so until their responses could be compared. It therefore also meant that the value of r solved above applied only to one sex at a time, but recognising of course that such value was intended to be only a guide to the author, because σ would probably change anyway as selection progressed. Finally, it was desirable that each replication be sampled for at least 30 pupae of each sex. This number would allow the author to draw accurate graphs of phenotypic distribution for any replication, and thereby to ascertain if selection was distorting the distribution sufficiently far from normal to warrant a scale transformation.

Consideration of all the above factors eventually led to the following experimental design. Selection in each population would involve six independent replications, each with its own Control, and each replication (and Control) would be sampled for 32 female and 32 male pupae per generation. However, this number of replications was raised to eight in the Light population for fear that unrelenting selection on such small animals may eventually cause some replications to die out. This did indeed happen, as described later, and on each occasion one of the extra replications replaced the one that had died out. Thus there were always six replications (but never more) analysed in each period.

SELECTION PROCEDURE: Six random samples, each of 32 female and 32 male pupae, were drawn from both base populations (Light and Heavy). The six were weighed and compared, to check that there were no significant differences among them. The variance analyses given below show that in the Heavy population there were indeed no significant differences in either female or male data; but in the

Light population significant differences appeared in both.

Source	df	MS	F expt	F .05(.01)	Result
	Heavy population		- female data		
Replications	5	0.094	1.71	2.26(3.11)	ns
Residual	186	0.055			
			- male data		
Replications	5	0.126	2.14	2.26(3.11)	ns
Residual	186	0.059			
	Light population		- female data		
Replications	7	0.266	4.52	2.05(2.73)	**
Residual	248	0.059			
			- male data		
Replications	7	0.230	3.54	2.05(2.73)	**
Residual	248	0.065			

The Light population was re-sampled and this second time there were no significant differences among the eight samples; they were therefore labelled as Replications 1 to 8.

In each generation selection was carried out for the heaviest (or lightest) four females and four males of the 32 pairs weighed. These eight parents were mass-mated, as there is apparently no trend for preferential mating in Tribolium (Shrode, 1960). From their resulting progeny a random 32 female and 32 male pupae were weighed and the heaviest (or lightest) four pairs of these in their turn were mass-mated to continue the selection programme.

As the Control for each replication reached its sixth generation, it was closed. A new Control was then extracted from the main selection line of the same replication.

This experiment was designed to enable a full analysis of selection response. A summary of its special features for this task is as follows -

(a) high replication, with all failures being replaced,

(b) short-term Controls, of a selection nature,

(c) response analysed over six-generation periods; and including the mean, the fluctuations, and the trends within this period,

(d) emphasis on the variation among replications,

(e) investigation into different ways of measuring the response.

However, computer analysis is not attempted in this thesis. The reason is simply that it was considered there would be insufficient time available to do an adequate job.

Fig 1 : SELECTION RESPONSE

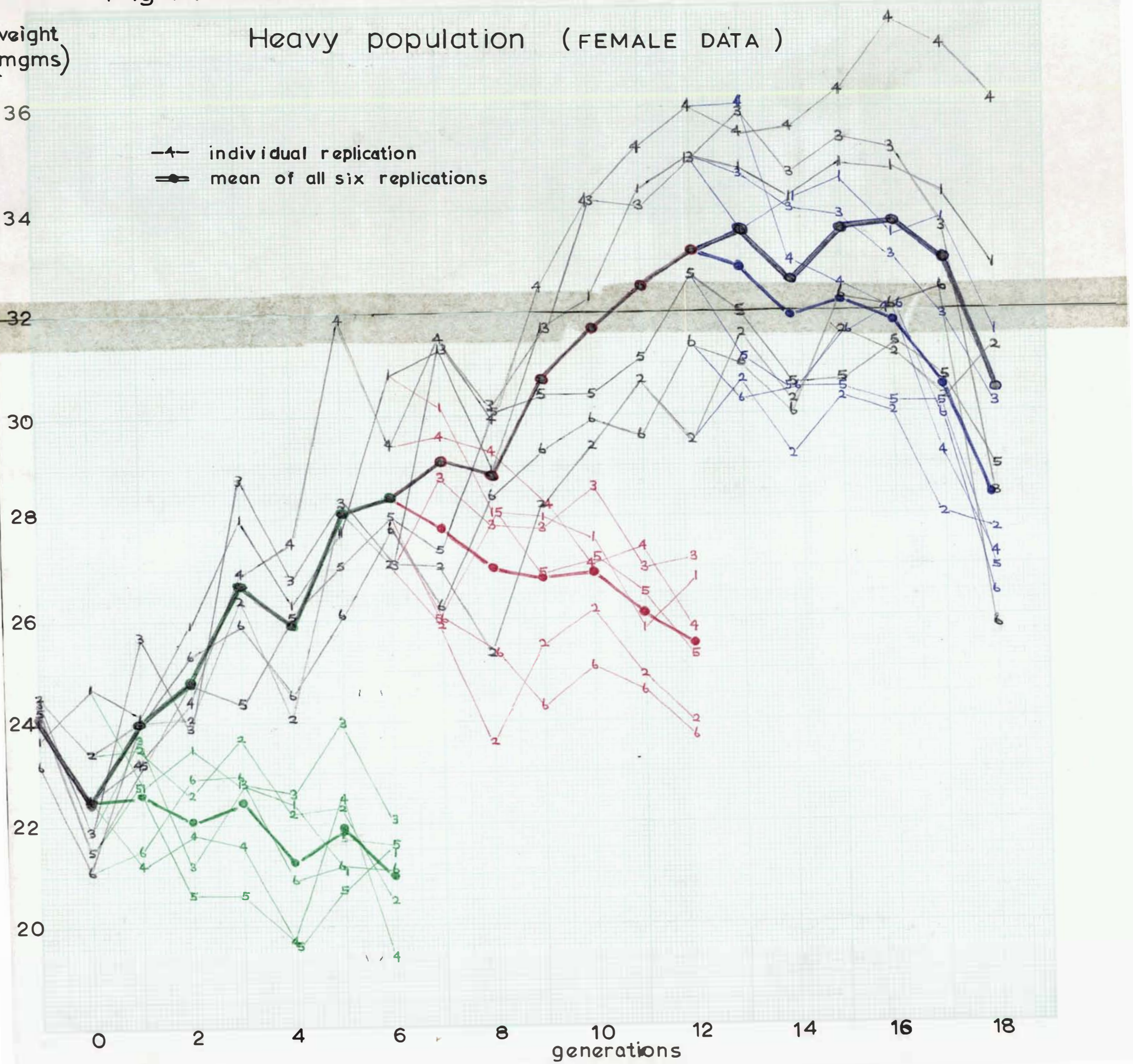


Fig 1 : continued

Heavy population (MALE DATA)

eight
ngms)

3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0

0 2 4 6 8 10 12 14 16 18

generations

Fig 1 : continued

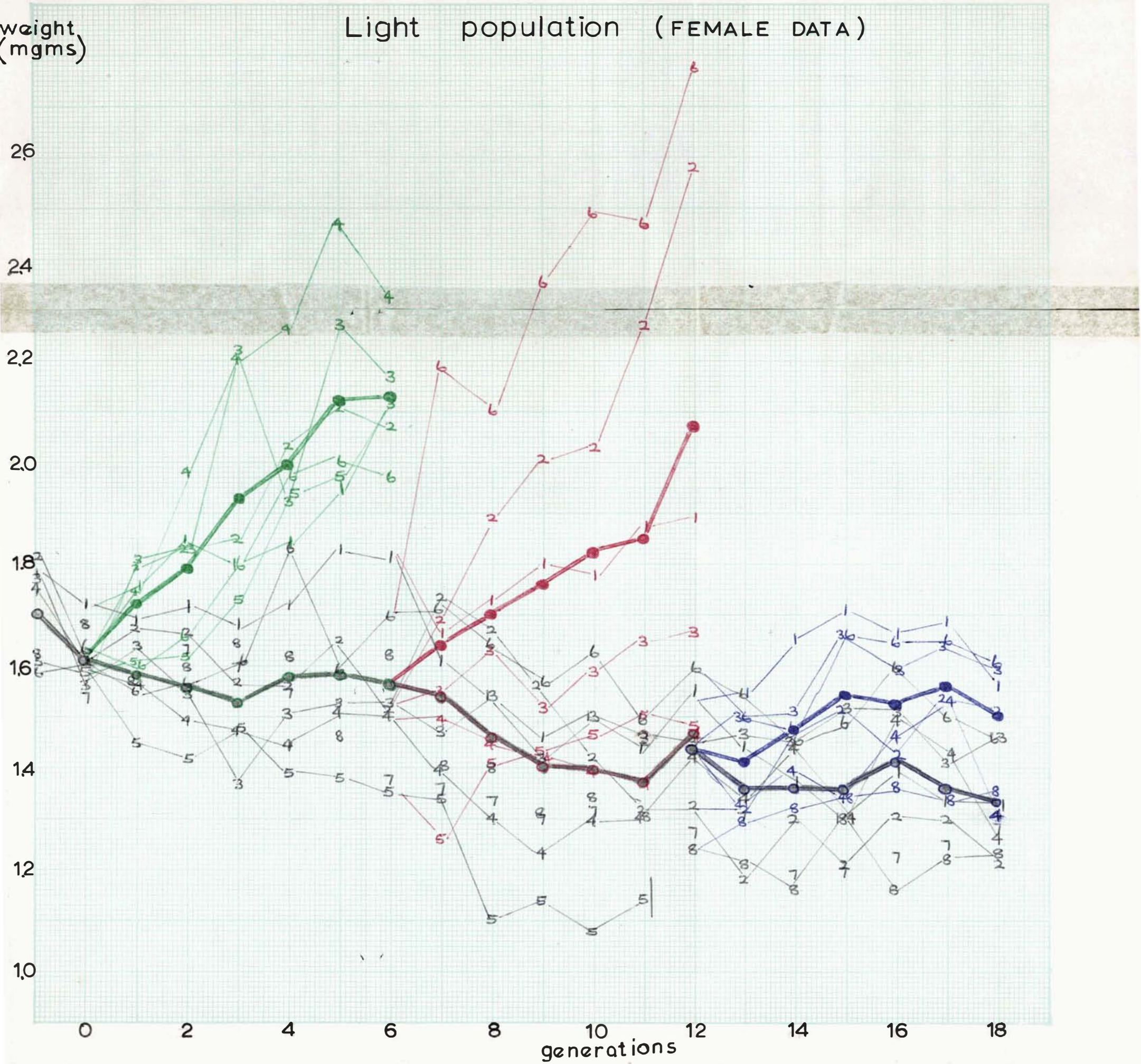
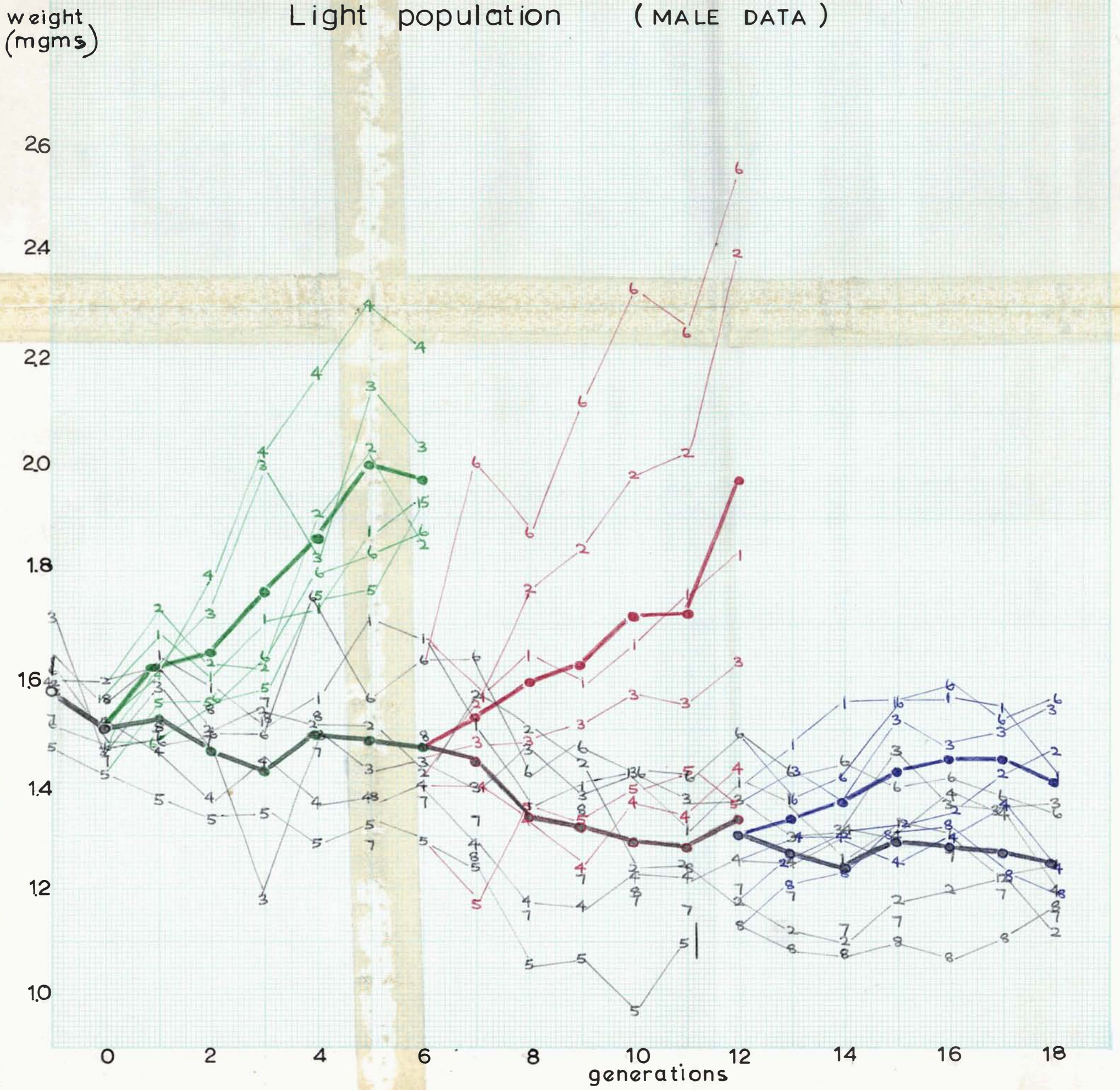


Fig 1 : continued

Light population (MALE DATA)



1.3 RESULTS:

In accordance with the plan described above, the first experiment of this thesis was set up and carried out. It lasted about twenty months and during this time only two small hitches occurred; firstly a thermostat broke down in the main incubator and all lines were killed by overheating. Reserve matings were used to restart the experiment and since fortunately the accident occurred in the first generation only time was lost and not results as well. Secondly it became urgent at about Generation 3 to shift all stocks from the D.S.I.R. incubators to Massey and there was not enough time to carefully calibrate the new incubators at Massey before the stocks were shifted into them. For a few subsequent days after shifting the incubator temperature was fluctuating quite markedly and this would probably have affected the pupal weights of that generation, both of the Controls and of the selected lines (Howe, 1956).

As the experiment progressed those lines in the Light population that were being selected for even smaller pupae began to lag slightly behind the Heavy lines. This situation arose because the small animals were apparently unable to produce progeny quite as fast as the larger ones could and so it took longer for the quota of 32 female and 32 male progeny to become available for measurement. In many cases there were never 32 such pairs at the correct stage of maturity available at any one time in a replication, which meant that the replication had to be sampled again a few days later. This made it difficult to determine exactly the generation interval in each population, particularly since there was a further confounding effect present - the fact that pressure of work occasionally ruled that a replication be sampled when the author was ready rather than when the replication itself was ready. However, it is interesting to trace even rough patterns of change in generation interval, and for purposes of the table given below the generation interval is calculated as

the number of days from when the first pupae are sampled and weighed in a replication until the first pupae of the next generation are sampled and weighed plus half the number of days that elapse before the last pupae required from this progeny generation are sampled and weighed. This correction factor caters for those Light replications in which sampling had to be spread out over several days, by assuming that on the average those pupae which were ultimately to be selected as parents were weighed halfway through the drawn-out sampling period. With this correction factor incorporated the following table lists the generation interval of the main upward and downward selection lines, averaged for each generation over all six replications. The Controls are not included in this table.

Table 2: Generation Interval (in days)

Generation	Heavy pop.	Light pop.
0 to 1	31	31
1 to 2	27	28
2 to 3	28	33
3 to 4	31	35
4 to 5	31	32
5 to 6	34	39
6 to 7	29	35
7 to 8	28	30
8 to 9	29	31
9 to 10	28	32
10 to 11	28	32
11 to 12	27	31
12 to 13	28	29
13 to 14	28	30
14 to 15	28	31
15 to 16	28	31
16 to 17	28	35
17 to 18	32	38
mean:	29	32

For every generation except the first one it took longer to produce the required number of progeny in the Light population than in the Heavy one. The reason for this is probably just an ability of big animals to lay eggs at a faster rate and to grow faster, rather than a question of say relative levels of inbreeding in the Heavy and Light populations. In neither population does the generation interval as measured this way show much pattern of change as selection proceeds, and this is surprising for it was expected that the selection character would be affected by either natural selection against the changing mean or by inbreeding depression. One of the Light replications died out at Generation 12 and another one was too weak to use for a diallel cross at Generation 18; since in neither case was there any evidence that accidents or disease caused the failures it seems likely that the continued pressures from artificial and natural selection were partly responsible.

(A) PRESENTATION OF BASIC DATA:

Although the collection of data on selection response is straightforward and does not involve difficult decisions, the presentation and analysis of this data requires considerable thought. This point arises because there are several alternative ways to present and analyse the data of an experiment and the choice of any one way would usually depend on the experimental aim. In the current case the author attempts to arrange, present and analyse the data in as many ways as possible, partly so that the different systems may be compared for value and partly because the different systems may give independent support to a hypothesis. In the first part of this Results section of the

experiment the raw data will be arranged and presented in different ways: in the second part it will be analysed statistically in more detail.

Mean weight: Not only is this parameter the logical first step in an analysis of results; it is usually also the most important one in a commercial sense. The mean weight data for this Tribolium selection experiment are presented for each generation and replication in Figure 1 (sexes separate) and Table 3 (sexes combined), and in each case the mean weight averaged over all replications is also given. The table of data is given here because it provides quicker reference if certain aspects of the data are to be studied in greater detail, but the graphs show up far more clearly the patterns of response and any anomalies in them. Some of these anomalies should be discussed before the patterns themselves are interpreted. Firstly consider the strong downward response at Generation 0 in the Heavy population, which coincided with the overheating accident that killed all main lines. The consequence of this accident that reserve matings had to be used instead, and that these reserves were lighter in weight and provided a smaller selection differential, probably helps to explain the lack of a positive response. Similarly the strongly downward response at Generation 3 in both selection line and Control of the Heavy population coincided with the transfer of stocks from D.S.I.R. to Massey, and it is known that the incubator climate was abnormal for a few days. It is recorded by Hove (1956) that such a climate change would affect pupae weights, though once again the Light population appears to have been less affected. Neither of these accidents affects the validity of results contained in the experiment for the Controls were of a nature able to cope with such accidents. Similarly the data are not greatly weakened by the death of Replication 5 after Generation 11 in the Light population, for a subsidiary replication (No.8) was immediately added to the analysis. The removal of No.5 from the analysis at

Table 3: Replication mean weights (mgms)

HEAVY POPULATION

	upward selection								downward Control						
Rep	1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen 0	2.38	2.30	2.14	2.16	2.12	2.04	2.19								
1	2.35	2.37	2.52	2.25	2.27	2.31	2.34	1	2.21	2.33	2.27	2.09	2.18	2.11	2.20
2	2.51	2.35	2.34	2.40	2.45	2.43	2.42	2	2.36	2.19	2.12	2.10	2.03	2.22	2.17
3	2.68	2.57	2.79	2.59	2.37	2.49	2.58	3	2.17	2.36	2.23	2.09	1.96	2.17	2.16
4	2.61	2.35	2.59	2.67	2.57	2.40	2.53	4	2.19	2.18	2.17	1.90	1.92	2.03	2.06
5	2.65	2.74	2.77	3.09	2.68	2.63	2.76	5	2.03	2.17	2.37	2.20	2.00	2.09	2.14
6	3.02	2.66	2.65	2.93	2.75	2.70	2.78	6	2.10	2.06	2.21	1.92	2.04	2.04	2.06
7	3.03	2.63	3.03	3.06	2.70	2.59	2.84	1	2.90	2.55	2.79	2.87	2.56	2.55	2.70
8	2.85	2.48	2.96	2.99	2.97	2.79	2.84	2	2.74	2.33	2.68	2.87	2.72	2.51	2.64
9	3.10	2.78	3.10	3.12	2.95	2.80	2.98	3	2.68	2.49	2.73	2.75	2.59	2.38	2.60
10	3.13	2.84	3.28	3.21	2.99	2.88	3.06	4	2.62	2.51	2.74	2.58	2.61	2.45	2.59
11	3.27	2.92	3.33	3.35	3.07	2.85	3.13	5	2.53	2.36	2.65	2.61	2.49	2.40	2.51
12	3.34	2.90	3.40	3.34	3.16	3.02	3.19	6	2.54	2.36	2.58	2.50	2.39	2.33	2.45
13	3.32	3.04	3.47	3.45	3.12	2.94	3.22	1	3.27	2.93	3.42	3.42	3.02	2.89	3.16
14	3.27	2.96	3.45	3.49	3.04	2.94	3.19	2	3.30	2.79	3.35	3.23	2.96	2.94	3.10
15	3.41	3.04	3.46	3.49	3.03	3.17	3.27	3	3.30	2.94	3.32	3.14	2.99	3.01	3.12
16	3.40	3.05	3.47	3.63	3.05	3.10	3.28	4	3.21	2.93	3.28	3.08	2.92	3.08	3.08
17	3.32	3.00	3.30	3.60	2.99	3.15	3.23	5	3.30	2.73	3.21	2.91	2.95	2.93	3.01
18	3.19	3.06	2.77	3.51	2.75	2.58	2.98	6	3.05	2.72	2.95	2.68	2.62	2.59	2.77

LIGHT POPULATION

	downward selection								upward Control						
Rep	1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen 0	1.64	1.60	1.51	1.55	1.51	1.55	1.56								
1	1.67	1.65	1.62	1.51	1.41	1.51	1.56	1	1.72	1.76	1.72	1.67	1.58	1.53	1.66
2	1.65	1.59	1.51	1.43	1.38	1.54	1.52	2	1.72	1.73	1.77	1.89	1.58	1.61	1.72
3	1.60	1.55	1.27	1.46	1.41	1.55	1.47	3	1.76	1.74	2.11	2.11	1.66	1.72	1.85
4	1.64	1.54	1.50	1.40	1.34	1.79	1.53	4	1.79	1.97	1.87	2.21	1.84	1.88	1.93
5	1.77	1.58	1.48	1.44	1.35	1.57	1.53	5	1.91	2.08	2.21	2.39	1.87	1.92	2.06
6	1.74	1.47	1.49	1.46	1.32	1.66	1.52	6	2.02	1.96	2.11	2.27	2.02	1.93	2.05
7	1.56	1.65	1.43	1.33	1.29	1.68	1.49	1	1.61	1.62	1.51	1.44	1.22	2.09	1.58
8	1.45	1.59	1.51	1.23	1.08	1.53	1.40	2	1.69	1.83	1.55	1.39	1.39	2.03	1.65
9	1.42	1.50	1.40	1.19	1.10	1.52	1.36	3	1.69	1.93	1.51	1.33	1.39	2.26	1.69
10	1.46	1.33	1.46	1.25	1.03	1.52	1.34	4	1.73	2.11	1.58	1.37	1.44	2.41	1.77
11	1.38	1.29	1.41	1.26	1.12	1.45	1.32	5	1.81	2.15	1.60	1.35	1.46	2.36	1.79
12	1.47	1.25	1.41	1.34	-	1.54	1.40	6	1.86	2.49	1.65	1.44	1.43	2.66	1.92
				Rep 8 →								Rep 8 →			
13	1.40	1.15	1.39	1.29	1.15	1.48	1.31	1	1.50	1.29	1.46	1.31	1.26	1.43	1.37
14	1.32	1.20	1.39	1.38	1.12	1.45	1.31	2	1.60	1.38	1.44	1.34	1.27	1.44	1.41
15	1.31	1.20	1.49	1.29	1.20	1.44	1.32	3	1.63	1.43	1.59	1.29	1.32	1.60	1.48
16	1.34	1.25	1.43	1.44	1.11	1.50	1.35	4	1.62	1.38	1.53	1.37	1.35	1.61	1.48
17	1.28	1.26	1.37	1.39	1.17	1.43	1.32	5	1.61	1.48	1.56	1.44	1.28	1.58	1.49
18	1.28	1.17	1.41	1.22	1.21	1.40	1.28	6	1.49	1.49	1.57	1.27	1.27	1.58	1.44

Generation 12, and also of its Control, has caused both means to rise sharply in that generation because that replication had been the lightest one by a large amount for some generations. No.8 was chosen to replace it because the difference between them was the least of those available, but the fact that their mean weights were still slightly unequal explains the presence of two overall means for Generation 12 in the graph.

Apart from these three anomalies the main patterns of response are clear. The Heavy population gave a strong and consistent response both in the main selection lines and in the Control lines until about Generation 12, and a weak response thereafter. The Light population followed a similar pattern except that downward progress was never as fast as in its upward Control, though it was usually just as regular. The standard genetical interpretation of these patterns would probably be that since selection response was so marked during the first two six-generation periods then there must have been considerable genetic variation still present in both Heavy and Light populations. But by Generation 12 this genetic variation appears to have been mostly used up, thereby permitting only a very slight subsequent response in main selection line and in Control. Thus it is evident that a graph such as Figure 1 can provide a lot of information, especially if the patterns are as clear as those obtained here, and that much more information on the performance of individual replications is equally present for the taking. However, such graphs can also depict phenomena without being able to explain them and in the present case the most noticeable of these phenomena is undoubtedly the asymmetry of response between main selection lines and Controls. Even without resorting to analysis it is clear that in both populations there is greater response to upward selection than to downward selection during the first two six-generation periods. There are several possible causes for this asymmetry, as discussed earlier, but such graphs as the ones given here are unable to

distinguish between these possibilities. The observation that this asymmetry seems to disappear during the third period tends to confuse the issue rather than to clarify it, and suggests that more than one source of asymmetry is involved. It remains to be seen if response parameters other than the population mean are of assistance in solving such problems.

Phenotypic variance: The values of phenotypic variance are given separately for each replication, generation, and sex in Table 4. A survey of this table shows that there is a large range of values within any six-generation period and a measure of this variability will be attempted later in the analysis of results. So great is this variability that even when all six replications are averaged within a generation, the generation means themselves show no clear patterns over the six stages of any period. Attempts to find such a pattern were therefore taken one stage further by concentrating on the nineteen successive generations of the main upward and downward selection lines, and by at the same time pooling the variance values for both sexes. This pooling was justified because the two sexes showed very parallel patterns of variance even though the female values were usually greater than those of the male. (This latter trend is likely to be a scale effect caused by the weight advantage of female over male). The resulting pattern of phenotypic variances obtained by averaging both sexes and all replications is probably not strictly valid in the sense that the component variances are not tested for homogeneity, but they should nevertheless indicate quite reliably if there is any trend of change in variance over nineteen generations. The resulting pattern is therefore given in Figure 2.

The Heavy population in Figure 2 shows a rise of variance until Generation 11 then shows a fall, but the Light population shows a fairly consistent fall throughout. Both patterns are blunted by strong fluctuations but they

Fig 2: RESPONSE IN OTHER PARAMETERS

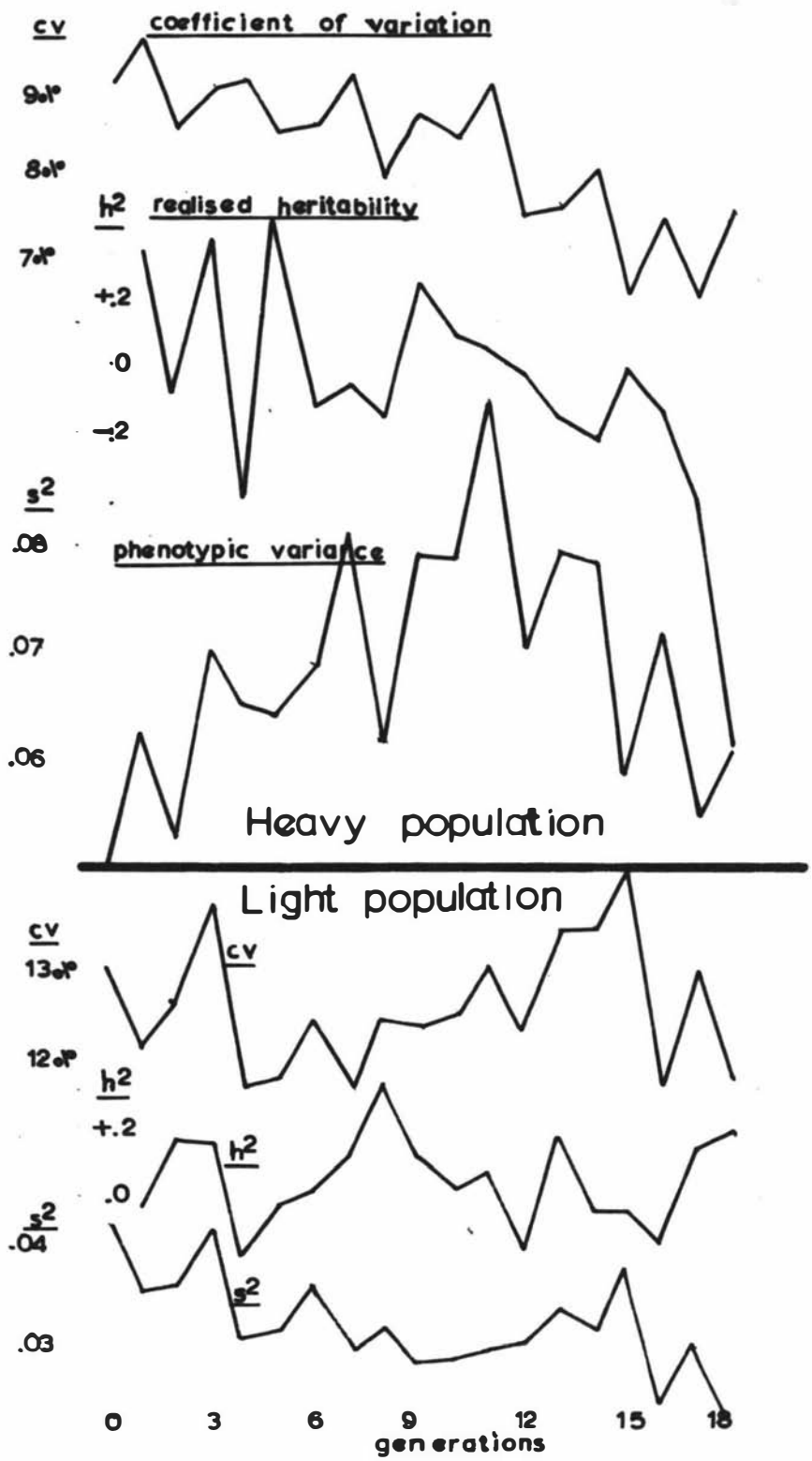


Table 4: Phenotypic variances a) Female data

HEAVY POPULATION

	upward selection								downward Control						
Rep	1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen 0	.053	.024	.033	.033	.050	.038	.039								
1	.050	.067	.046	.052	.066	.040	.053	1	.053	.043	.073	.044	.042	.060	.052
2	.037	.026	.035	.074	.026	.042	.040	2	.044	.033	.044	.075	.064	.035	.049
3	.071	.061	.039	.066	.037	.070	.057	3	.065	.061	.038	.026	.037	.078	.051
4	.043	.055	.064	.058	.056	.043	.053	4	.047	.065	.061	.026	.044	.035	.046
5	.042	.051	.038	.050	.071	.060	.052	5	.045	.037	.050	.036	.023	.046	.039
6	.032	.055	.077	.064	.063	.063	.059	6	.086	.059	.028	.041	.070	.094	.063
7	.082	.107	.088	.038	.042	.058	.069	1	.068	.051	.047	.076	.043	.085	.062
8	.061	.071	.023	.036	.050	.083	.054	2	.043	.049	.046	.034	.084	.059	.053
9	.068	.070	.063	.035	.047	.077	.060	3	.067	.083	.050	.081	.102	.057	.073
10	.053	.077	.038	.076	.059	.118	.070	4	.049	.025	.043	.118	.058	.053	.058
11	.031	.086	.038	.087	.059	.136	.073	5	.072	.077	.069	.117	.083	.034	.075
12	.018	.064	.020	.070	.031	.074	.046	6	.089	.079	.062	.074	.088	.055	.074
13	.054	.066	.018	.081	.028	.096	.057	1	.038	.074	.030	.033	.040	.091	.051
14	.049	.090	.066	.043	.048	.079	.063	2	.057	.074	.045	.044	.028	.125	.062
15	.052	.051	.029	.061	.018	.058	.045	3	.033	.049	.070	.105	.040	.028	.054
16	.047	.051	.035	.038	.067	.088	.054	4	.051	.023	.092	.089	.036	.042	.055
17	.043	.046	.079	.041	.036	.030	.046	5	.046	.063	.065	.049	.025	.022	.045
18	.067	.061	.064	.040	.053	.044	.055	6	.073	.046	.059	.054	.073	.077	.064

LIGHT POPULATION

	downward selection								upward Control						
Rep	1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen 0	.042	.031	.041	.054	.065	.050	.047								
1	.032	.027	.032	.040	.048	.027	.034	1	.028	.040	.025	.035	.043	.043	.036
2	.043	.032	.037	.042	.041	.032	.038	2	.036	.065	.042	.044	.056	.038	.047
3	.029	.035	.061	.034	.033	.071	.044	3	.043	.048	.035	.051	.048	.040	.044
4	.040	.029	.026	.026	.032	.035	.031	4	.048	.026	.039	.032	.054	.029	.038
5	.038	.042	.034	.024	.037	.024	.033	5	.052	.039	.027	.030	.046	.043	.040
6	.042	.032	.034	.040	.037	.037	.037	6	.046	.044	.041	.026	.049	.048	.042
7	.048	.035	.017	.037	.024	.040	.034	1	.023	.030	.023	.038	.025	.094	.039
8	.029	.067	.033	.036	.016	.041	.037	2	.018	.054	.046	.050	.012	.078	.043
9	.026	.039	.028	.024	.022	.031	.028	3	.031	.066	.033	.057	.027	.074	.048
10	.027	.026	.018	.041	.018	.041	.029	4	.020	.065	.028	.025	.033	.071	.040
11	.033	.032	.027	.032	.019	.053	.033	5	.022	.083	.015	.016	.014	.084	.039
12	.025	.044	.022	.039	-	.040	.034	6	.047	.066	.039	.031	.031	.036	.042
				Rep 8 →								Rep 8 →			
13	.037	.025	.047	.056	.023	.062	.042	1	.035	.018	.032	.054	.028	.041	.035
14	.028	.022	.038	.043	.032	.037	.033	2	.031	.045	.044	.048	.029	.049	.041
15	.017	.031	.024	.041	.047	.070	.038	3	.036	.049	.046	.032	.024	.049	.039
16	.016	.021	.032	.027	.030	.037	.027	4	.032	.022	.067	.038	.017	.035	.035
17	.018	.018	.020	.034	.058	.039	.031	5	.027	.048	.022	.022	.039	.033	.032
18	.021	.021	.035	.023	.022	.030	.025	6	.050	.026	.031	.018	.047	.056	.038

Table 4: (continued) b) Male data

HEAVY POPULATION

	upward selection								downward Control						
Rep	1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen 0	.053	.038	.046	.027	.049	.041	.042								
1	.099	.044	.020	.051	.054	.041	.052	1	.057	.045	.075	.042	.061	.039	.053
2	.033	.061	.035	.063	.041	.045	.046	2	.050	.036	.044	.037	.045	.074	.048
3	.080	.054	.060	.085	.036	.060	.062	3	.055	.055	.038	.028	.038	.077	.048
4	.033	.098	.051	.082	.035	.044	.057	4	.026	.062	.036	.038	.037	.032	.038
5	.039	.060	.047	.058	.039	.096	.056	5	.046	.053	.044	.044	.037	.072	.049
6	.060	.058	.057	.040	.078	.059	.059	6	.042	.034	.036	.036	.053	.056	.043
7	.081	.119	.072	.067	.042	.055	.073	1	.129	.096	.111	.054	.063	.034	.081
8	.061	.034	.038	.056	.056	.056	.050	2	.040	.046	.040	.061	.096	.065	.058
9	.102	.082	.073	.067	.052	.092	.078	3	.059	.071	.054	.107	.092	.067	.075
10	.050	.086	.077	.067	.034	.089	.067	4	.090	.051	.110	.052	.102	.054	.070
11	.102	.116	.063	.112	.063	.103	.093	5	.067	.086	.055	.058	.068	.052	.064
12	.060	.059	.086	.139	.039	.064	.075	6	.086	.078	.100	.050	.084	.074	.079
13	.089	.066	.098	.093	.052	.088	.081	1	.042	.064	.041	.123	.083	.079	.072
14	.117	.100	.023	.056	.033	.109	.073	2	.064	.067	.059	.049	.062	.108	.068
15	.042	.068	.017	.108	.042	.037	.052	3	.103	.037	.049	.136	.046	.073	.074
16	.045	.062	.046	.093	.063	.106	.069	4	.047	.061	.101	.068	.076	.088	.073
17	.066	.022	.052	.045	.042	.038	.044	5	.045	.036	.039	.075	.036	.024	.043
18	.058	.023	.045	.056	.038	.063	.047	6	.101	.049	.039	.043	.048	.040	.053

LIGHT POPULATION

	downward selection								upward Control						
Rep	1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen 0	.041	.032	.038	.040	.040	.036	.038								
1	.030	.057	.021	.042	.035	.042	.038	1	.025	.032	.034	.027	.028	.027	.029
2	.033	.035	.035	.034	.044	.030	.035	2	.042	.046	.043	.071	.030	.032	.044
3	.045	.041	.053	.037	.030	.035	.040	3	.033	.025	.031	.065	.033	.034	.037
4	.033	.038	.022	.030	.028	.038	.031	4	.030	.037	.037	.049	.034	.039	.038
5	.028	.040	.023	.039	.029	.033	.032	5	.051	.030	.036	.023	.048	.037	.038
6	.029	.020	.029	.032	.035	.072	.036	6	.041	.046	.032	.048	.059	.043	.045
7	.027	.035	.017	.026	.019	.037	.027	1	.036	.034	.019	.026	.019	.073	.034
8	.028	.038	.033	.018	.017	.033	.028	2	.031	.027	.034	.028	.012	.060	.032
9	.022	.045	.019	.023	.020	.053	.030	3	.033	.117	.024	.020	.014	.045	.042
10	.026	.028	.033	.034	.013	.046	.030	4	.029	.029	.015	.017	.021	.103	.036
11	.024	.019	.031	.035	.027	.039	.029	5	.025	.045	.033	.018	.016	.112	.042
12	.024	.015	.018	.046	-	.038	.028	6	.029	.095	.027	.035	.024	.058	.045
				Rep 8 →								Rep 8 →			
13	.044	.015	.014	.025	.015	.051	.027	1	.028	.020	.016	.034	.018	.025	.024
14	.037	.020	.031	.032	.024	.042	.031	2	.039	.017	.023	.033	.018	.022	.025
15	.029	.015	.035	.042	.049	.052	.037	3	.022	.033	.025	.014	.021	.053	.028
16	.013	.023	.022	.032	.012	.041	.024	4	.024	.018	.017	.020	.017	.034	.022
17	.012	.033	.012	.044	.045	.038	.031	5	.033	.024	.027	.030	.039	.035	.031
18	.020	.016	.011	.019	.025	.038	.021	6	.026	.017	.033	.027	.039	.049	.032

are still clear enough to be strikingly similar to the corresponding patterns for mean weight given in Figure 1. This similarity is very important from the viewpoint of analysing results for it indicates that the two parameters, mean weight and phenotypic variance, might be related in the sense that they may both be affected by some other factor such as scale. If such was the case it would be invalid to treat mean weight and phenotypic variance as independent interpreters of genetic behaviour. What is needed therefore is some other measure of variability that is free from the scale effect which appears to bias the mean weight and variance results. Such a possible alternative measure is called the coefficient of variation and this alternative is now discussed.

Coefficient of variation: This parameter is a measure of population variability that differs from the previous measure by taking into account the population mean weight. It is based on the conception that a given level of variability about a light population mean weight is biologically more significant than the same absolute level of variability about a heavy population mean. The statistical definition of coefficient of variation is symbolised as s/μ and its value in the present context is in being able to cancel scale effects that are due to the logarithmic nature of growth. That is, if the phenotypic variance of a population was seen to rise during selection and this rise was in fact a scale effect caused by the logarithmic growth process in association with a rising mean weight, then conversion of the data to coefficient of variation would portray the variability as being constant in value. Because there is adequate reason to believe that scale effects would be present and indeed are present, the current data on phenotypic variability has therefore been converted to coefficient of variation and is presented in Table 5.

Table 5: Coefficients of variation (%) a) Female data

HEAVY POPULATION

Rep	upward selection 1	2	3	4	5	6	mean		downward Control 1	2	3	4	5	6	mean
Gen 0	9.3	6.7	8.3	8.1	10.4	9.2	8.7								
1	9.2	10.8	8.4	9.8	11.1	8.4	9.6	1	10.1	8.8	11.4	9.9	8.9	11.4	10.1
2	7.5	6.8	7.8	11.2	6.5	8.1	8.0	2	8.9	8.1	9.9	12.5	12.3	8.2	10.0
3	9.5	9.3	6.8	9.6	7.9	10.2	8.9	3	11.2	10.4	8.6	7.5	9.4	12.2	9.9
4	7.8	9.7	9.4	8.8	9.1	8.5	8.9	4	9.7	11.5	10.9	8.1	10.7	9.0	10.0
5	7.3	8.0	6.9	7.0	9.8	9.4	8.1	5	10.2	8.6	9.3	8.4	7.3	10.2	9.0
6	5.8	8.6	10.3	8.6	9.0	9.0	8.5	6	13.7	11.9	7.6	10.5	12.3	14.6	11.8
7	9.1	12.1	9.5	6.2	7.5	9.2	8.9	1	8.6	8.7	7.6	9.3	7.9	11.2	8.9
8	8.5	10.5	5.1	6.3	7.4	10.2	8.0	2	7.4	9.3	7.7	6.3	10.3	9.6	8.4
9	8.2	9.4	7.9	5.8	7.1	9.5	8.0	3	9.2	11.3	8.0	10.1	11.9	9.8	10.1
10	7.2	9.5	5.7	8.1	8.0	11.5	8.3	4	8.0	6.1	7.3	12.7	8.8	9.2	8.7
11	5.1	9.5	5.7	8.3	7.8	12.4	8.1	5	10.4	11.1	9.8	12.5	10.9	7.5	10.4
12	3.8	8.6	4.1	7.4	5.3	8.7	6.3	6	11.2	11.7	9.1	10.5	11.7	9.9	10.7
13	6.7	8.1	3.8	8.0	5.2	10.0	7.0	1	5.8	8.9	5.0	5.0	6.4	10.0	6.8
14	6.5	9.9	7.4	5.8	7.2	9.3	7.7	2	7.0	9.3	6.3	6.3	5.5	3.7	6.3
15	6.5	7.1	4.8	6.8	4.4	7.5	6.2	3	5.2	7.3	7.8	10.0	6.6	5.3	7.0
16	6.2	7.2	5.3	5.2	8.2	9.2	6.9	4	6.7	5.1	9.1	9.3	6.3	6.4	7.1
17	6.1	7.1	8.4	5.5	6.2	5.4	6.4	5	6.3	9.0	8.0	7.6	5.3	5.0	6.9
18	7.9	7.9	8.9	5.5	8.2	8.1	7.8	6	8.6	7.7	8.1	8.6	10.0	10.5	8.9

LIGHT POPULATION

Rep	downward selection 1	2	3	4	5	6	mean		upward Control 1	2	3	4	5	6	mean
Gen 0	11.9	10.9	12.9	14.6	15.8	13.7	13.3								
1	10.5	9.9	10.8	12.8	15.2	10.6	11.6	1	9.6	11.2	8.7	10.7	12.8	13.0	11.0
2	12.1	10.8	12.5	13.7	14.3	11.4	12.5	2	10.2	14.0	11.1	10.6	14.6	11.8	12.0
3	10.2	12.0	18.0	12.5	12.3	16.5	13.6	3	11.5	11.9	8.4	10.3	12.7	11.1	11.0
4	11.7	10.8	10.6	11.3	12.8	10.1	11.2	4	11.9	8.0	10.3	7.9	12.1	8.7	9.8
5	10.6	12.4	12.0	10.1	14.0	9.8	11.5	5	11.7	9.4	7.3	7.1	10.9	10.4	9.5
6	11.2	11.7	12.0	13.1	14.2	11.4	12.3	6	10.0	10.1	9.3	7.0	10.4	11.1	9.7
7	13.6	10.8	9.0	13.9	11.5	11.7	11.7	1	9.2	10.2	9.8	13.1	12.6	14.1	11.5
8	11.0	15.5	11.8	14.5	11.4	12.3	12.8	2	7.8	12.3	13.2	15.6	7.7	13.3	11.6
9	11.0	12.6	11.8	12.6	13.0	11.3	12.0	3	9.7	12.8	12.0	16.7	11.6	11.6	12.4
10	11.0	11.5	9.0	15.6	12.3	12.5	12.0	4	7.9	11.4	10.6	11.4	12.2	10.8	10.7
11	12.7	13.6	11.2	13.7	12.1	15.4	13.1	5	8.0	12.7	7.4	9.2	7.7	11.7	9.5
12	10.1	15.9	10.2	14.0	-	12.6	12.6	6	11.5	10.0	11.8	12.0	10.3	6.8	10.4
				Rep 8 →								Rep 8 →			
13	13.3	13.5	14.8	17.6	12.6	16.2	14.7	1	12.2	10.1	12.0	17.6	12.9	13.3	13.0
14	12.2	11.5	13.4	14.3	15.3	13.1	13.3	2	10.7	14.4	14.0	15.8	13.0	14.9	13.8
15	10.0	14.4	10.2	15.5	16.7	17.8	14.1	3	11.0	14.6	13.0	13.4	11.5	13.3	12.8
16	9.0	10.9	11.8	11.0	15.1	12.1	11.6	4	10.7	10.4	16.3	13.4	9.6	11.4	12.0
17	10.2	10.2	10.1	13.0	19.5	13.2	12.7	5	9.8	14.3	9.0	9.8	14.9	11.1	11.5
18	10.8	12.0	12.8	12.0	12.1	11.8	11.9	6	14.3	10.6	11.0	10.2	16.0	14.9	12.8

Table 4: Phenotypic variances a) Female data

HEAVY POPULATION

Rep	upward selection 1	2	3	4	5	6	mean		downward Control 1	2	3	4	5	6	mean
Gen 0	.053	.024	.033	.033	.050	.038	.039								
1	.050	.067	.046	.052	.066	.040	.053	1	.053	.043	.073	.044	.042	.060	.052
2	.037	.026	.035	.074	.026	.042	.040	2	.044	.033	.044	.075	.064	.035	.049
3	.071	.061	.039	.066	.037	.070	.057	3	.065	.061	.038	.026	.037	.078	.051
4	.043	.055	.064	.058	.056	.043	.053	4	.047	.065	.061	.026	.044	.035	.046
5	.042	.051	.038	.050	.071	.060	.052	5	.045	.037	.050	.036	.023	.046	.039
6	.032	.055	.077	.064	.063	.063	.059	6	.086	.059	.028	.041	.070	.094	.063
7	.082	.107	.088	.038	.042	.058	.069	1	.068	.051	.047	.076	.043	.085	.062
8	.061	.071	.023	.036	.050	.083	.054	2	.043	.049	.046	.034	.084	.059	.053
9	.068	.070	.063	.035	.047	.077	.060	3	.067	.083	.050	.081	.102	.057	.073
10	.053	.077	.038	.076	.059	.118	.070	4	.049	.025	.043	.118	.058	.053	.058
11	.031	.086	.038	.087	.059	.136	.073	5	.072	.077	.069	.117	.083	.034	.075
12	.018	.064	.020	.070	.031	.074	.046	6	.089	.079	.062	.074	.088	.055	.074
13	.054	.066	.018	.081	.028	.096	.057	1	.038	.074	.030	.033	.040	.091	.051
14	.049	.090	.066	.043	.048	.079	.063	2	.057	.074	.045	.044	.028	.125	.062
15	.052	.051	.029	.061	.018	.058	.045	3	.033	.049	.070	.105	.040	.028	.054
16	.047	.051	.035	.038	.067	.088	.054	4	.051	.023	.092	.089	.036	.042	.055
17	.043	.046	.079	.041	.036	.030	.046	5	.046	.063	.065	.049	.025	.022	.045
18	.067	.061	.064	.040	.053	.044	.055	6	.073	.046	.059	.054	.073	.077	.064

LIGHT POPULATION

Rep	downward selection 1	2	3	4	5	6	mean		upward Control 1	2	3	4	5	6	mean
Gen 0	.042	.031	.041	.054	.065	.050	.047								
1	.032	.027	.032	.040	.048	.027	.034	1	.028	.040	.025	.035	.043	.043	.036
2	.043	.032	.037	.042	.041	.032	.038	2	.036	.065	.042	.044	.056	.038	.047
3	.029	.035	.061	.034	.033	.071	.044	3	.043	.048	.035	.051	.048	.040	.044
4	.040	.029	.026	.026	.032	.035	.031	4	.048	.026	.039	.032	.054	.029	.038
5	.038	.042	.034	.024	.037	.024	.033	5	.052	.039	.027	.030	.046	.043	.040
6	.042	.032	.034	.040	.037	.037	.037	6	.046	.044	.041	.026	.049	.048	.042
7	.048	.035	.017	.037	.024	.040	.034	1	.023	.030	.023	.038	.025	.094	.039
8	.029	.067	.033	.036	.016	.041	.037	2	.018	.054	.046	.050	.012	.078	.043
9	.026	.039	.028	.024	.022	.031	.028	3	.031	.066	.033	.057	.027	.074	.048
10	.027	.026	.018	.041	.018	.041	.029	4	.020	.065	.028	.025	.033	.071	.040
11	.033	.032	.027	.032	.019	.053	.033	5	.022	.083	.015	.016	.014	.084	.039
12	.025	.044	.022	.039	-	.040	.034	6	.047	.066	.039	.031	.031	.036	.042
				Rep 8 →								Rep 8 →			
13	.037	.025	.047	.056	.023	.062	.042	1	.035	.018	.032	.054	.028	.041	.035
14	.028	.022	.038	.043	.032	.037	.033	2	.031	.045	.044	.048	.029	.049	.041
15	.017	.031	.024	.041	.047	.070	.038	3	.036	.049	.046	.032	.024	.049	.039
16	.016	.021	.032	.027	.030	.037	.027	4	.032	.022	.067	.038	.017	.035	.035
17	.018	.018	.020	.034	.058	.039	.031	5	.027	.048	.022	.022	.039	.033	.032
18	.021	.021	.035	.023	.022	.030	.025	6	.050	.026	.031	.018	.047	.056	.038

Table 4: (continued) b) Male data

HEAVY POPULATION

		upward selection								downward Control						
Rep		1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen	0	.053	.038	.046	.027	.049	.041	.042								
	1	.099	.044	.020	.051	.054	.041	.052	1	.057	.045	.075	.042	.061	.039	.053
	2	.033	.061	.035	.063	.041	.045	.046	2	.050	.036	.044	.037	.045	.074	.048
	3	.080	.054	.060	.085	.036	.060	.062	3	.055	.055	.038	.028	.038	.077	.048
	4	.033	.098	.051	.082	.035	.044	.057	4	.026	.062	.036	.038	.037	.032	.038
	5	.039	.060	.047	.058	.039	.096	.056	5	.046	.053	.044	.044	.037	.072	.049
	6	.060	.058	.057	.040	.078	.059	.059	6	.042	.034	.036	.036	.053	.056	.043
	7	.081	.119	.072	.067	.042	.055	.073	1	.129	.096	.111	.054	.063	.034	.081
	8	.061	.034	.038	.056	.056	.056	.050	2	.040	.046	.040	.061	.096	.065	.058
	9	.102	.082	.073	.067	.052	.092	.078	3	.059	.071	.054	.107	.092	.067	.075
	10	.050	.086	.077	.067	.034	.089	.067	4	.050	.051	.110	.052	.102	.054	.070
	11	.102	.116	.063	.112	.063	.103	.093	5	.067	.086	.055	.058	.068	.052	.064
	12	.060	.059	.086	.139	.039	.064	.075	6	.086	.078	.100	.050	.084	.074	.079
	13	.089	.066	.098	.093	.052	.088	.081	1	.042	.064	.041	.123	.083	.079	.072
	14	.117	.100	.023	.056	.033	.109	.073	2	.064	.067	.059	.049	.062	.108	.068
	15	.042	.068	.017	.108	.042	.037	.052	3	.103	.037	.049	.136	.046	.073	.074
	16	.045	.062	.046	.093	.063	.106	.069	4	.047	.061	.101	.068	.076	.088	.073
	17	.066	.022	.052	.045	.042	.038	.044	5	.045	.036	.039	.075	.036	.024	.043
	18	.058	.023	.045	.056	.038	.063	.047	6	.101	.049	.039	.043	.048	.040	.053

LIGHT POPULATION

		downward selection								upward Control						
Rep		1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen	0	.041	.032	.038	.040	.040	.036	.038								
	1	.030	.057	.021	.042	.035	.042	.038	1	.025	.032	.034	.027	.028	.027	.029
	2	.033	.035	.035	.034	.044	.030	.035	2	.042	.046	.043	.071	.030	.032	.044
	3	.045	.041	.053	.037	.030	.035	.040	3	.033	.025	.031	.065	.033	.034	.037
	4	.033	.038	.022	.030	.028	.038	.031	4	.030	.037	.037	.049	.034	.039	.038
	5	.028	.040	.023	.039	.029	.033	.032	5	.051	.030	.036	.023	.048	.037	.038
	6	.029	.020	.029	.032	.035	.072	.036	6	.041	.046	.032	.048	.059	.043	.045
	7	.027	.035	.017	.026	.019	.037	.027	1	.036	.034	.019	.026	.019	.073	.034
	8	.028	.038	.033	.018	.017	.033	.028	2	.031	.027	.034	.028	.012	.060	.032
	9	.022	.045	.019	.023	.020	.053	.030	3	.033	.117	.024	.020	.014	.045	.042
	10	.026	.028	.033	.034	.013	.046	.030	4	.029	.029	.015	.017	.021	.103	.036
	11	.024	.019	.031	.035	.027	.039	.029	5	.025	.045	.033	.018	.016	.112	.042
	12	.024	.015	.018	.046	-	.038	.028	6	.029	.095	.027	.035	.024	.058	.045
					Rep 8 →								Rep 8 →			
	13	.044	.015	.014	.025	.015	.051	.027	1	.028	.020	.016	.034	.018	.025	.024
	14	.037	.020	.031	.032	.024	.042	.031	2	.039	.017	.023	.033	.018	.022	.025
	15	.029	.015	.035	.042	.049	.052	.037	3	.022	.033	.025	.014	.021	.053	.028
	16	.013	.023	.022	.032	.012	.041	.024	4	.024	.018	.017	.020	.017	.034	.022
	17	.012	.033	.012	.044	.045	.038	.031	5	.033	.024	.027	.030	.039	.035	.031
	18	.020	.016	.011	.019	.025	.038	.021	6	.026	.017	.033	.027	.039	.049	.032

Table 5: (Continued) b) Male data

HEAVY POPULATION

Rep	upward selection 1	2	3	4	5	6	mean		downward Control 1	2	3	4	5	6	mean
Gen 0	10.0	8.7	10.3	8.0	10.6	10.3	9.6								
1	13.7	8.9	5.8	10.3	10.6	9.1	9.7	1	11.1	9.3	12.6	10.0	11.8	9.5	10.7
2	7.4	10.8	8.2	10.6	8.4	9.1	9.1	2	9.5	8.9	9.9	9.5	10.6	12.6	10.2
3	11.0	9.3	9.1	11.6	8.2	10.2	9.9	3	11.3	10.0	8.9	8.2	10.5	13.6	10.4
4	7.0	13.7	9.0	11.1	7.5	8.9	9.5	4	7.5	11.6	9.2	10.6	10.2	9.1	9.7
5	7.8	9.2	7.9	8.1	7.5	11.7	8.7	5	11.0	10.8	8.9	9.8	10.0	12.9	10.6
6	8.3	9.2	9.2	6.8	10.4	9.3	8.9	6	9.9	8.9	8.6	10.0	12.0	12.0	10.2
7	9.7	13.5	9.2	8.7	7.7	9.2	9.7	1	12.9	12.4	12.3	8.3	10.0	7.4	10.5
8	8.8	7.6	6.8	7.9	8.0	8.6	8.0	2	7.5	9.3	7.7	8.8	11.9	10.2	9.2
9	10.6	10.5	8.9	8.6	8.0	11.3	9.6	3	9.5	11.0	8.7	12.2	12.1	11.1	10.8
10	7.4	10.7	8.8	8.6	6.3	10.7	8.8	4	9.0	9.4	12.7	9.2	12.7	9.8	10.5
11	10.3	12.3	7.7	10.5	8.3	11.8	10.1	5	10.5	13.2	9.0	9.7	11.2	9.8	10.6
12	7.7	8.5	8.9	12.1	6.4	8.7	8.7	6	12.2	12.0	12.9	9.3	12.8	12.0	11.9
13	9.4	8.8	4.6	9.1	7.5	10.7	8.3	1	6.4	9.1	6.0	10.8	9.8	10.2	8.7
14	11.0	11.0	4.5	6.9	6.0	11.5	8.5	2	8.0	9.7	7.3	7.0	8.7	11.6	8.7
15	6.1	8.9	3.9	9.8	6.8	6.2	7.0	3	10.2	6.7	6.8	12.2	7.3	9.5	8.8
16	6.3	8.4	6.3	8.8	8.4	10.9	8.2	4	7.1	8.7	9.7	8.8	9.8	10.1	9.0
17	8.0	5.0	7.0	6.0	7.0	6.4	6.6	5	6.6	7.1	6.2	9.5	6.6	5.3	6.9
18	7.8	5.1	7.9	6.9	7.2	9.7	7.4	6	10.8	8.3	6.9	7.8	8.6	7.8	8.4

LIGHT POPULATION

Rep	downward selection 1	2	3	4	5	6	mean		upward Control 1	2	3	4	5	6	mean
Gen 0	13.0	11.2	13.4	13.2	14.1	12.9	13.0								
1	10.5	14.7	9.2	14.0	13.7	13.8	12.7	1	9.5	10.3	11.4	10.2	10.9	11.1	10.6
2	11.5	12.4	12.8	13.5	15.9	11.5	12.9	2	12.7	13.2	12.1	14.9	11.2	11.5	12.6
3	14.0	13.2	19.4	13.3	12.9	12.5	14.2	3	10.7	9.8	8.8	12.6	11.4	11.3	10.8
4	11.6	12.8	9.9	12.8	13.1	11.2	11.9	4	10.0	10.1	10.6	10.2	10.6	11.0	10.4
5	9.8	13.2	10.5	14.4	13.0	11.7	12.1	5	12.1	8.4	8.8	6.5	12.5	10.5	9.8
6	10.3	9.9	11.8	12.8	14.5	16.5	12.6	6	10.6	11.6	8.8	9.8	12.7	11.0	10.7
7	10.9	12.0	9.5	12.6	11.0	11.7	11.3	1	12.2	11.9	9.3	11.5	11.7	13.5	11.7
8	12.2	12.9	12.3	11.4	12.4	12.7	12.3	2	10.7	9.3	12.4	12.6	8.1	12.5	10.9
9	10.6	14.8	10.0	13.0	13.3	15.6	12.9	3	11.5	18.6	10.2	11.4	8.8	9.8	11.7
10	11.4	13.4	12.7	15.1	11.7	15.3	13.3	4	10.2	8.6	7.8	9.7	10.3	13.8	10.1
11	11.8	11.1	12.8	15.4	15.0	14.1	13.4	5	9.0	10.5	11.7	10.1	8.8	14.9	10.8
12	11.2	10.2	9.7	17.2	-	13.1	12.3	6	9.3	12.8	10.1	13.2	12.8	9.4	11.3
				Rep 8 →								Rep 8 →			
13	15.4	11.0	9.1	12.7	11.4	15.8	12.6	1	11.4	11.3	9.0	14.2	11.2	11.5	11.4
14	15.3	12.7	13.4	13.6	14.4	14.2	13.9	2	12.8	10.1	11.0	14.1	10.9	10.5	11.6
15	13.0	10.2	12.9	15.9	20.2	16.4	14.8	3	9.5	13.6	10.5	9.6	11.0	14.8	11.5
16	8.6	12.5	10.9	13.1	10.4	14.4	11.6	4	10.0	10.0	9.0	11.0	9.8	11.7	10.2
17	8.9	14.9	8.3	15.6	19.1	14.2	13.5	5	11.8	11.0	10.9	12.9	16.1	12.2	12.5
18	11.4	11.2	7.8	11.5	13.5	14.6	11.7	6	11.5	8.9	11.7	13.1	16.6	14.2	12.7

Table 5 shows once again a considerable range of values between individual replications within a generation and between successive generations within a replication, so the former have been averaged to give more clarity of pattern. When this is done it is noticeable that the generation mean coefficients of variation for females are mostly less than those for males, and also that in the Heavy population the generation means for upward selection are nearly always less than those for the downward Control. Almost as frequently in the Light population are the upward Control means less than those of the downward line. The striking point about all three comparisons is that in each one the member with greater mean weight almost always has the smaller coefficient of variation, for example females are heavier than males but their coefficient of variation is less. Thus there is apparently an inverse correlation between mean weight and coefficient of variation, which suggests that conversion of the data to the latter parameter has over-compensated for the scale effect. The pattern for coefficient of variation averaged over all replications and both sexes is given in Figure 2 for the main upward and downward selection lines. In the Heavy population there is a clear tendency for coefficient of variation to fall in successive generations as the mean weight rises. The Light population does not show such a clear tendency to rise as the mean weight falls, except over the middle generations. Its subsequent fall in later generations may be due to the removal of Replication 5, containing the smallest animals, from all analyses beyond Generation 12. Notwithstanding this particular pattern, the overall data suggests that scale effects are not sufficiently strong to warrant measuring the phenotypic variability in terms of coefficient of variation. It is not sensible to speculate too deeply on why this should be so but two possible reasons stand out. Firstly the scaling effect caused by logarithmic growth processes acting on a given amount of genetic variability will be reduced if the genetic variability itself is reduced, and this latter

is theoretically expected to happen under continued selection. Secondly the effect of scale on phenotypic variability may be confounded with the effect of an opposing force such as natural selection trying to reduce the variability.

Whatever may be the exact cause(s) of such relationships it thus seems evident that neither of the two parameters, variance or coefficient of variation, are free of a relationship with mean weight in the present experiment. In the absence of better alternatives it also seems likely that this relationship is due to scale and natural selection. However, one parameter which is claimed as being almost unaffected by either scale or natural selection (Falconer, 1961) is called realised heritability and it is discussed below as an alternative system of interpreting a selection response.

Realised heritability: This parameter is calculated from the formula h^2 = selection response/selection differential. The reason it is recommended for such analyses of response as the present one is that any factor such as scale or natural selection which affects the selection differential is likely to affect the selection response in the same direction and by a near-proportional amount. That is, numerator and denominator of the above fraction will be affected to about the same relative extent, leaving the fraction itself almost unchanged. Realised heritability is therefore a parameter likely to give a clearer idea of the true genetic response to selection than are other parameters, and so the appropriate data are given in Table 6.

It is very difficult to see any patterns among the data of Table 6 even when the replications are averaged in each generation. It is not discounted that some patterns may indeed be present but over any of the six-generation periods considered they are masked by the very large fluctuations occurring between generations. As was done for the two earlier parameters discussed,

Table 6: Realised heritabilities a) Female data

HEAVY POPULATION

		upward selection								downward Control						
Rep		1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen	1	-.26	.23	1.15	.28	.68	1.00	.51	1	.76	-.04	-.73	.46	-.27	-.11	.01
	2	.62	.03	-.60	.34	.35	.41	.19	2	-.21	.27	.71	-.22	.79	-.37	.16
	3	.75	.79	1.48	.58	-.17	.22	.61	3	.25	-.38	-.50	.05	.00	.00	-.10
	4	-.49	-.70	-.59	.19	.53	-.39	-.24	4	.10	.37	.05	.65	.32	.43	.32
	5	.38	1.00	.38	1.26	.29	.40	.62	5	.38	-.02	-.36	-1.40	-.35	-.07	-.30
	6	.88	-.29	-.35	-.69	.29	.39	.04	6	-.12	.64	.47	1.03	-.30	.00	.29
	7	.21	-.03	1.02	.55	-.17	-.47	.19	1	.19	.31	-.39	-.05	.53	.42	.17
	8	-.71	-.36	-.21	-.64	.97	.63	-.05	2	.84	.68	.24	.06	-.60	.16	.23
	9	.68	.69	.65	.93	.10	.24	.55	3	.03	-.53	.03	.33	.24	.24	.06
	10	.16	.36	1.00	.57	.00	.15	.37	4	.10	-.17	-.19	.19	-.06	-.19	-.05
	11	.62	.36	-.04	.36	.23	-.07	.24	5	.50	.43	.41	-.06	.17	.10	.26
	12	.33	-.30	.31	.20	.55	.33	.24	6	-.28	.21	-.04	.30	.23	.25	.11
	13	-.10	.66	.41	-.18	-.30	-.14	.06	1	.61	-.25	.13	-.02	.59	.20	.21
	14	-.19	-.36	.67	.03	-.52	-.22	-.10	2	-.19	.33	.23	.91	.15	-.04	.23
	15	.21	.35	.20	.27	.03	.57	.27	3	-.09	-.24	.03	.12	.00	-.17	-.06
	16	-.04	-.11	-.08	.48	.30	-.06	.08	4	.36	.06	.14	.07	.09	-.11	.10
	17	-.15	-.24	-.61	-.22	-.16	.09	-.21	5	-.07	.71	.23	.57	.00	.50	.32
	18	-.48	.31	-1.21	-.44	-1.12	-2.54	-.91	6	.61	.06	.39	.48	1.10	1.21	.64

LIGHT POPULATION

		downward selection								upward Control						
Rep		1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
	1	.09	-.30	-.28	.11	.33	.19	.02	1	.14	.63	.73	.38	.00	-.10	.30
	2	-.10	.00	.33	.23	.09	-.08	.08	2	.31	.14	.08	.77	.04	.16	.25
	3	.10	.37	.54	.06	-.20	-.13	.12	3	-.14	.05	1.62	.65	.33	.48	.50
	4	-.12	.00	-.35	.10	.35	-.71	-.12	4	.13	.51	-1.20	.17	.59	.46	.11
	5	-.37	-.32	-.08	-.41	.04	.86	-.05	5	.29	.36	1.06	.77	.12	.08	.45
	6	.03	.37	.00	-.08	.11	-.45	.00	6	.56	-.17	-.36	-.67	.50	-.07	-.03
	7	.78	-.78	.21	.42	.07	-.07	.10	1	-.46	.51	.06	-.17	-.30	1.75	.23
	8	.21	.18	-.41	.31	1.00	.21	.25	2	.29	.61	.30	-.16	.67	-.22	.25
	9	.33	.26	.25	.25	-.27	.19	.17	3	.19	.32	-.32	-.05	.06	.59	.13
	10	-.15	.48	-.36	-.33	.35	-.19	-.03	4	-.07	.43	.21	-.08	.17	.31	.16
	11	.24	.40	.25	-.03	-.33	.42	.16	5	.37	.09	.29	-.08	.07	-.03	.12
	12	-.38	-.04	.00	-.48	-	-.34	-.25	6	.09	.58	.13	.45	-.10	.86	.33
					Rep 8 →								Rep 8 →			
	13	.48	.54	-.08	.25	.16	.16	.25	1	-.04	-.02	.25	-.40	.29	-.39	-.05
	14	.18	-.67	.03	-.53	.24	.19	-.09	2	.42	.75	.04	.23	.07	-.05	.24
	15	.38	.39	-.19	.52	-.74	-.11	.04	3	.28	.13	.42	-.19	.07	.46	.19
	16	-.50	-.62	.04	-.80	.50	-.27	-.27	4	-.21	-.36	-.19	.36	.11	-.07	-.06
	17	.32	.09	.34	.23	-.35	.42	.17	5	.07	.48	.08	.24	-.20	.00	.11
	18	.00	.45	-.24	1.38	-.03	.08	.27	6	-.50	-.09	-.21	-1.04	.09	-.15	-.32

Table 6: (Continued) b) Male data

HEAVY POPULATION

		upward selection								downward Control						
Rep		1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen	1	-.04	.37	1.15	.37	.38	.87	.52	1	.75	-.18	-.26	.11	-.03	-.31	.01
	2	.37	-.23	-1.20	.49	.54	.31	.05	2	-.52	.42	.14	.10	.22	-.23	.02
	3	.46	.69	1.29	.43	-.37	.21	.45	3	.74	-.66	-.16	.00	.48	.21	.10
	4	.03	-.70	-.54	.20	.68	-.08	-.07	4	-.22	.60	.31	.76	-.06	.24	.27
	5	-.22	.70	.51	.91	.46	1.11	.58	5	.75	.04	-.77	-1.00	-.18	-.38	-.26
	6	1.50	-.15	-.47	-.19	.18	-.07	.13	6	-.31	.14	.41	.61	.06	.19	.18
	7	-.17	-.15	.95	.17	-.09	-.21	.08	1	.38	.27	-.29	.44	.42	.32	.26
	8	-.29	-.22	-.09	.03	1.04	.53	.17	2	.29	.47	.21	-.07	-.24	.00	.11
	9	.51	.88	.52	.07	-.19	-.25	.26	3	.37	-.48	-.34	.28	.31	.45	.10
	10	.06	.07	.27	.00	.27	.22	.15	4	.16	.06	.15	.39	-.02	-.17	.09
	11	.19	.05	.45	.45	.37	-.09	.24	5	.03	.49	.05	-.11	.32	.12	.15
	12	.22	.20	.14	-.25	.14	.37	.14	6	.19	-.29	.32	.28	.19	.13	.14
	13	-.04	.16	.12	.44	-.09	-.41	.03	1	.00	.17	-.10	-.34	.34	.36	.07
	14	-.14	-.10	.20	.22	-.04	.25	.06	2	.03	.29	.22	.12	.11	-.19	.10
	15	.49	.12	-.17	-.26	-.12	.53	.10	3	.06	-.45	.13	.36	-.12	-.06	-.01
	16	.00	.21	.12	.33	-.06	-.41	.03	4	.10	.03	-.02	.11	.25	-.20	.04
	17	-.63	.00	-.42	.00	-.17	.19	-.17	5	-.33	.37	.05	.13	-.13	.12	.03
	18	-.33	.04	-1.81	-.23	-.61	-1.96	-.82	6	.67	.00	.95	.51	.89	1.07	.68

LIGHT POPULATION

		downward selection								upward Control						
Rep		1	2	3	4	5	6	mean		1	2	3	4	5	6	mean
Gen	1	-.28	-.10	-.42	.14	.14	-.03	-.09	1	.50	.50	.47	.35	.46	.00	.38
	2	.25	.30	.59	.23	.12	-.06	.24	2	-.36	-.42	.33	.95	.00	.36	.14
	3	.21	-.08	1.17	-.23	-.03	.04	.18	3	.37	-.03	.74	.50	.09	.27	.32
	4	-.11	.06	-.94	.28	.22	-.86	-.22	4	.06	1.26	-.72	.43	.59	.55	.36
	5	-.52	.00	.25	-.06	-.15	.49	.00	5	.48	.45	1.00	.56	.08	.15	.45
	6	.14	.31	-.04	-.06	.12	-.21	.04	6	.18	-.83	-.41	-.37	.47	.15	-.13
	7	.50	-.68	.18	.46	.36	-.03	.13	1	-.44	.56	.15	.00	-.43	.71	.09
	8	.52	.23	-.47	.52	.95	.65	.40	2	.28	.73	.00	-.21	.76	-.09	.24
	9	-.16	.20	.45	.06	-.06	-.20	.05	3	-.18	.28	.12	-.37	-.06	.46	.04
	10	-.12	.62	-.22	-.32	.50	.20	.11	4	.29	.25	.25	.48	.20	.57	.34
	11	.45	-.04	.17	.00	-.76	.00	-.03	5	.30	.15	-.10	-.14	.11	-.18	.02
	12	-.36	.35	.00	-.12	-	-.31	-.09	6	.33	.95	.36	.45	-.25	.58	.40
					Rep 8 →								Rep 8 →			
	13	.14	.29	.35	.00	.25	.26	.21	1	.28	.50	.25	.17	.30	-.42	.18
	14	.31	.13	-.10	-.39	.05	-.06	-.01	2	.32	.23	-.25	-.04	.08	.17	.08
	15	-.25	-.32	-.56	.12	-.09	.17	-.15	3	.00	.14	.54	-.16	.37	.54	.24
	16	.07	-.25	.32	-.28	.10	-.09	-.02	4	.07	.07	-.20	.24	.10	.08	.06
	17	.44	-.13	.08	.11	-.23	.14	.07	5	-.13	.28	.17	.23	-.50	-.27	-.04
	18	-.17	.40	-.18	.62	-.21	.12	.10	6	-.50	.16	.18	-.41	-.07	.14	-.08

the realised heritability values are averaged for both sexes and graphed over all nineteen generations in Figure 2, the main selection lines only being given. The resulting graphs show a slight then strong fall in the Heavy population but no apparent trend at all in the Light population. These are two different patterns of heritability response and so they require different interpretations of genetic behaviour in the two populations. In the first population there is a tendency for heritability to decline and this signifies that there is a decline in the amount of additive genetic variance present. In the second population there is no such tendency and this suggests that the genetic variability itself is not being reduced noticeably. In both cases there is at least a qualitative agreement between the interpretations of genetic behaviour from the heritability data and from the earlier data on mean weight, at least when the Control lines are ignored. But the ability to test for a more quantitative agreement is spoiled by the very large fluctuations that are present and since such fluctuations are undoubtedly due partly to changes in environment, it is clear that these Control lines should not be so ignored.

It appears fairly simple to sum up the contribution that graphs can make towards the analysis of a selection response. Undoubtedly they are the best way to show if there are any trends present in the response and if such trends are strong enough it is possible that the genetic behaviour during selection can be interpreted quite safely from graphs alone. But where such trends are slight or are masked by fluctuations the graphs are usually unable to give more than a qualitative assessment of the genetic behaviour,

and this is the case in the present experiment. Even when the values of all six replications and both sexes were averaged the resulting graphs for the parameters of phenotypic variance, coefficient of variation, and realised heritability fluctuated too greatly in the six-generation Control li s to show any distinct trend. This situation calls for a more thorough statistical analysis of the data, partly in the hope that trends invisible to the eye may be detected by statistics and partly to measure the variability between replications (which were pooled for the graphs). The way in which this analysis was approached in the present experiment will now be described.

(B) METHODS OF ANALYSING DATA :

It would seem wise to place this approach into perspective immediately. The system devised and given below was adapted to the present experiment and may not apply in all details to experiments with a different type of Control, but it is considered that the general concept of analysing trends and fluctuations as well as the mean response, is worth considering for all experiments. The actual benefits from doing this will probably depend on the actual sizes of fluctuations and trends in a given experiment and in the present case these fluctuations are big enough to warrant measurement separately. Hence the analytical approach used in the present experiment is to treat each six-generation period separately and in each period to measure the mean response, the fluctuations and the trends over all six generations.

Average rate of response: Perhaps the standard method of measuring the average progress of a response curve is by fitting a range of known mathematical curves to it, and so determining which of the models best fits the observed curve. By an iterative process the mathematical models can be made increasingly refined, until the deviation of actual graph points from the theoretical points of the model falls below an accepted level of significance. The mathematical formula behind this theoretical curve which best fits the actual data may then be used both to interpret the genetic behaviour of the population and to measure its selection response. For example, in 1950 Cavalli analysed some data of Dobzhansky by successively improving the fit of his mathematical model to the data using an iterative method of maximum likelihood. The same process has been carried out in more detail and with more meaning recently by James (1965). Some selection responses reported by four earlier authors were re-analysed by James and in each case he tested the data for goodness-of-fit

against three mathematical models - straight line, parabola, and exponential curve. In all four cases the parabola gave a better fit than did the straight line, mostly significantly better: and the exponential curve consistently gave a better fit than did the parabola. The interpretation of this result is that the selection response was falling off with time at an exponential rate. The formula of the curve will provide a measure of this rate and hence of the selection response itself. The fact that a straight line curve did not fit the data as well as the others means that the selection response cannot be considered as linear, and that therefore means the standard procedure of measuring mean response by linear regression must not be done without serious thought first. (The further fact that all three of the curves applied to the data did not give a very close fit does not lessen the merit of this approach. Its weakness is in the labour involved, because a close fit would usually result only after the testing and modifying of many complex genetic models and without the aid of a computer such labour is rarely worthwhile). This serious thought appears to have been given by Falconer (1953), who was apparently one of the early authors to measure mean response from the coefficient of linear regression. Unfortunately his approach seems to have been adopted without much further thought by later authors until now it is apparently regarded as fairly standard procedure (Fig. 5 of Kojima and Kelleher, 1963). Falconer argued that the estimation of mean response from linear regression is a good method because it minimises the squared deviations of individual generation means from the regression line. In this least-squares sense it thus provides a better fit to the data than does estimation of the mean response from simply dividing the total response by the number of generations. But two assumptions are implicit in such use of linear regression to estimate mean response: first, that there be reason to expect a truly linear response and second that there be no reason to expect a correlation between successive generation means. Falconer was not prepared to accept

these assumptions for his experiment as be argued that the genetic theory behind such selection process predicts that the response will in fact be non-linear and that successive means will be correlated, for example by genetic drift. However, the observed response over his thirteen generations was fairly linear, and he calculated that drift was probably unimportant in his particular case also. The approach in this present experiment is also not to automatically consider regression analysis as being valid but instead to test for this validity and to measure also how far the estimate of mean response from linear regression differs from the estimate obtained by a less refined method.

The first point to make in this comparison is that although the response curve of Figure 1 averaged over all replications could be regarded in most periods as being linear, those curves of the individual replications fluctuate too much to be so regarded. Moreover these individual replications should be analysed separately because of the heterogeneity between them. (The high amount of fluctuation and the low degrees of freedom in each replication make it not worthwhile trying to fit regression curves of higher order than linear).

In the analysis that follows, the selection response of the Heavy population during the first period of six generations has been used to compare methods of assessing response. The mean response for each replication separately is measured by fitting a straight regression line to the data and then the results are tested for the two assumptions given above - that there is no significant departure of the response curves from linearity, and that there is no correlation between the means of separate generations. In this analysis the Controls are treated separately from the main selection lines. The generation means are given for both cases in Table 7.

Table 7: Mean response from regression analysis (Gens 0 to 6 in Heavy population)

a) selection upwards

	Rep 1	2	3	4	5	6	Total(G)
Gen 0	2.38	2.30	2.14	2.16	2.12	2.04	13.14
1	2.35	2.37	2.52	2.25	2.27	2.31	14.07
2	2.51	2.35	2.34	2.40	2.45	2.43	14.48
3	2.68	2.57	2.79	2.59	2.37	2.49	15.49
4	2.61	2.35	2.59	2.67	2.57	2.40	15.19
5	2.65	2.74	2.77	3.09	2.68	2.63	16.56
6	3.02	2.66	2.65	2.93	2.75	2.70	16.71
Total (R)	18.20	17.34	17.80	18.09	17.21	17.00	105.64(T)
mean (μ)	2.60	2.48	2.54	2.58	2.46	2.43	
regr. coeff. (b)	.094	.065	.081	.151	.101	.092	

b) selection downwards (Control)

	Rep 1	2	3	4	5	6	Total(G)
Gen 0	2.38	2.30	2.14	2.16	2.12	2.04	13.14
1	2.21	2.33	2.27	2.09	2.18	2.11	13.19
2	2.36	2.19	2.12	2.10	2.03	2.22	13.02
3	2.17	2.36	2.23	2.09	1.96	2.17	12.98
4	2.19	2.18	2.17	1.90	1.92	2.03	12.39
5	2.03	2.17	2.37	2.20	2.00	2.09	12.86
6	2.10	2.06	2.21	1.92	2.04	2.04	12.37
Total (R)	15.44	15.59	15.51	14.46	14.25	14.70	89.95(T)
mean	2.21	2.23	2.22	2.07	2.04	2.10	
regr. coeff.	-.049	-.037	.016	-.025	-.025	-.008	

TEST FOR NON-LINEARITY: The regression coefficients given above are calculated from the standard formula. It is possible to test if these regression coefficients are significantly different from zero but in the present context it is more important to test if they adequately describe the relation between mean weight and generation number in the sense that variability of the former is caused mostly by change in the latter. This test of adequacy, or linearity, of the regression coefficient thus is resolved by a variance analysis in which it is tested if the amount of variability that remains between generations after removing the fraction due to the regression slope of response, is still a significant amount. The structure of this variance analysis is given below (adapted from Mather, 1949a) using the first table above as a numerical example.

Table 8a: Deviations from linear regression

a) selection upwards

Source of Variance	df	MS	F	P
Replications	5	0.034		
Regression	1	1.600	11.43	.001
Deviation from Regr.	5	0.018	1.29	.20
Residual	30	0.014		
Total	41			

b) selection downwards (Control)

Source of Variance	df	MS	F	P
Replications	5	0.050		
Regression	1	0.080	9.60	.01
Deviation from Regr.	5	0.008	0.96	-
Residual	30	0.0083		
Total	41			

In both of the variance tables there is a highly significant regression effect and in neither table is there any significant deviation of values from the regression lines. This result indicates that despite the apparently large fluctuations of individual replications there is no evidence of non-linearity in them; on this basis linear regression provides a valid and adequate measure of the mean response. The same conclusions may of course not be obtained if this same analysis was carried out on the data from the remaining two periods and from the Light population. However, a superficial survey of the response curves in Figure 1 suggests that the same results would indeed be obtained except for Period 3 of the Heavy population, and in this latter case any significant non-linearity would undoubtedly result from the environmental disturbance in Generation 18. All graphs of response divergence between main selection line and Control are quite linear (Figure 3).

TEST FOR CORRELATION BETWEEN GENERATIONS: The aim of this section is to determine if the guinea-pig data presented above obey the requirement that deviations from a regression line are uncorrelated with one another. As Mood and Graybill (1963) point out, the model for linear regression may be written as

$$y_i = \alpha + \beta x_i + e_i$$

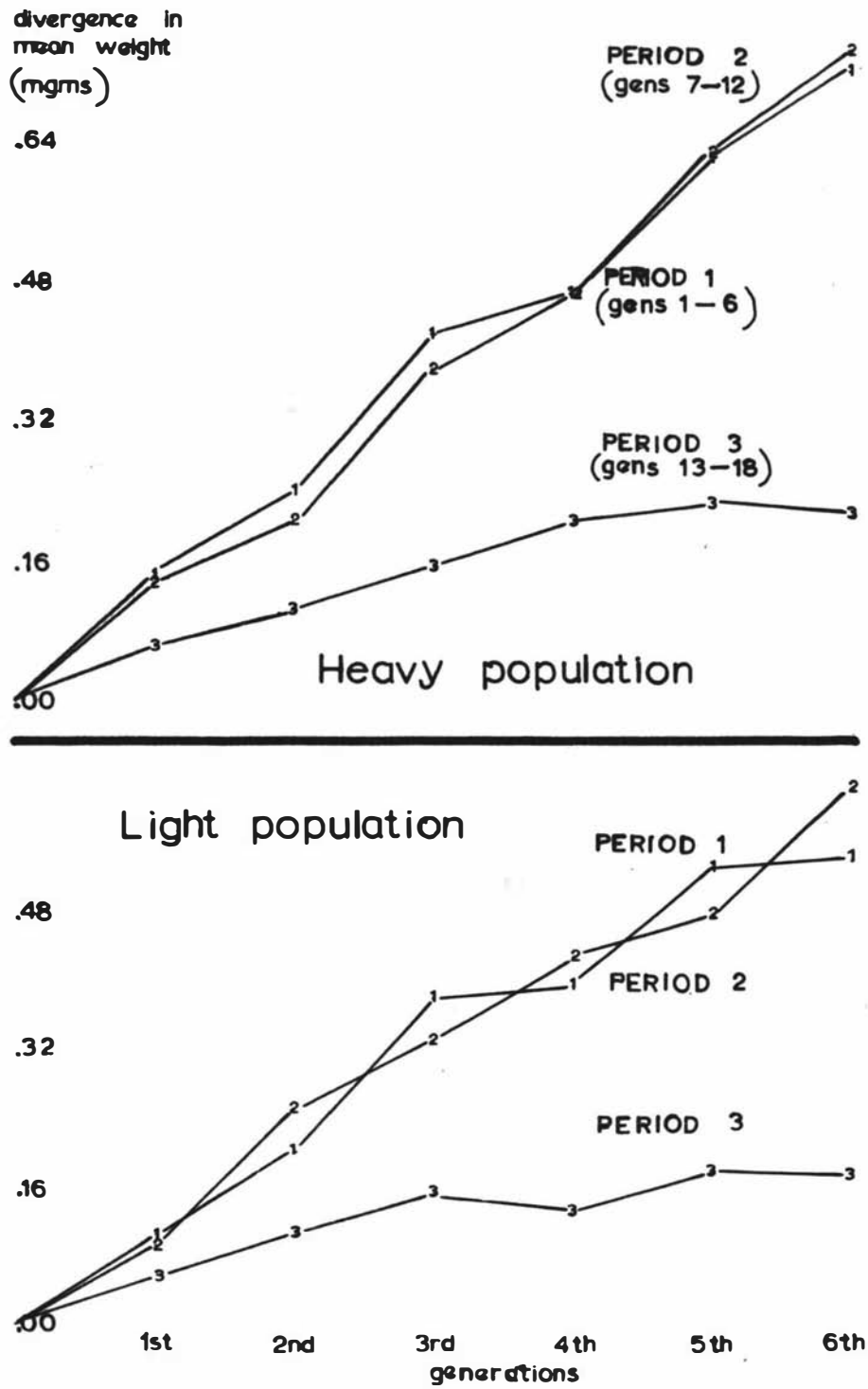
where x_i are known variables

y_i are measurable variables

α and β are constants determined from the data

e_i are unobservable random variables that are uncorrelated with one another and have a mean of zero, variance σ^2. Only under these conditions are the Best Linear Unbiased Estimators of α and β obtained

Fig 3: DIVERGENCE OF RESPONSE



by least squares analysis of the data, as was done by regression above. If the e_i are instead correlated with one another then this traditional least squares system of estimating the regression coefficient, and thence the mean selection response, is invalid in the pure sense. These e_i are estimated from the above data as the deviations of individual generation means from the regression line. There is obviously a very large number of ways in which these deviation values can be combined in order to test for correlations between them: however, it was considered that if such correlation did indeed exist it was most likely to be present between the means of adjacent generations. Accordingly the test required amounts to a measurement of the serial correlation between deviations taken two at a time, that is the correlation coefficient between the pairs of deviations from Generations 0 and 1, 1 and 2, 2 and 3, and so on.

The first procedure for this test of correlation is to calculate the actual deviation of each generation mean from its regression line, and the data for this calculation is contained in Table 7. Each deviation is obtained by the formula

$$dev_i = y_i - [\mu + b(i-3)]$$

where i = number of Generation

The resulting sets of deviations are given below in Table 8b, for the same six generations in the Heavy population.

The required correlation coefficients between adjacent pairs of deviations are calculated from the standard formula

$$r = \frac{\Sigma (x_i - \bar{x}_A)(x_{i+1} - \bar{x}_B)}{\sqrt{\Sigma (x_i - \bar{x}_A)^2 \cdot \Sigma (x_{i+1} - \bar{x}_B)^2}}$$

Table 8b: Correlations between deviations

a) selection upwards

	Rep 1	2	3	4	5	6
Gen 0	.06	.01	-.16	.04	-.04	-.11
1	-.06	.02	.14	-.03	.01	.06
2	.00	-.07	-.12	-.03	.09	.09
3	.08	.09	.25	.01	-.09	.06
4	-.08	-.19	-.03	-.06	.01	-.12
5	-.14	.13	.07	.21	.02	.02
6	.14	-.01	-.13	-.11	-.01	-.01
corr. coeff(r)	-.443	-.761	-.777	-.675	-.483	-.161

b) selection downwards (Control)

Gen 0	.02	-.04	-.03	.02	.00	-.08
1	-.10	.03	.08	-.03	.09	-.01
2	.10	-.08	-.08	.01	-.04	.11
3	-.04	.13	.01	.02	-.08	.07
4	.03	-.01	-.07	-.15	-.09	-.06
5	-.08	.01	.12	.18	.01	.01
6	.04	-.06	-.06	-.08	.08	-.04
corr. coeff.	-.799	-.601	-.753	-.755	+.234	+.071

where i = number of Generation concerned

$\bar{x}_A$ = mean deviation of Generations 0 to 5 inclusive

$\bar{x}_B$ = mean deviation of Generations 1 to 6 inclusive

x_i = individual deviation.

The correlation coefficients estimated by the above formula are contained

in Table 8b. On five occasions out of the twelve these coefficients are significantly negative, as determined from the serial correlation tables of Anderson (1942). The existence of this negative correlation warrants discussion. It has already been stated that the model for linear regression is

$$y_i = \alpha + \beta x_i + e_i$$

and that the e_i are uncorrelated, that is $E(e_i e_j) = 0$. When this zero expectation is shown in a given case to be not realised then the validity of this general linear model must be doubted. However there does not appear to be any discussion on this point in Graybill's book (1961) on linear models nor in Mood and Graybill (1963) and so there may be no easy way of estimating the amount of bias that this correlation causes. The method by which the correlation was calculated, namely from the deviations of adjacent means taken two at a time, interprets a negative coefficient as meaning that any large positive deviation at one generation tends to be followed or preceded by a negative or small positive deviation. It is worthwhile speculating on how this correlation could arise. Since no evidence of non-linear response was observed in the above example it may be tentatively accepted that the regression line traces the true selection response of the population, and that deviations from this line are caused by weighing errors or sampling errors and so should be wholly random. But the fact that it was always the sample itself, and not the larger population, from which parents were chosen to continue the selection line, suggests a possible solution. Suppose that by chance a sample was taken which was atypically heavier than the population it was meant to represent. Under the normal assumptions of selection theory, namely that additive genetic variance is evenly spread through the population, the parents selected from this atypical sample should produce exactly the same heritability of response as would any other individuals used as parents. Under this theoret-

ical assumption therefore atypical samples in one generation would not bias the mean of the succeeding generation so as to produce a negative correlation. However, if this theoretical assumption is not correct and instead all individuals in a population do not have the same breeding value, such a bias is possible. Suppose that in an upward selection line the heaviest individuals of a sample have the lowest breeding values and the lightest individuals have the highest breeding values. Then if an atypically heavy sample is taken from the population and atypically heavy individuals are chosen to be parents, the heritability of response from these parents will be atypically small. This gives exactly the negative correlation effect observed. The original sampling error would have led to a larger-than-expected response in the first instance but the lowered heritability of the chosen parents automatically would have led to a smaller-than-expected response in the succeeding generation. Consider also the case in which a sample contains atypically small animals. If these small animals have higher breeding values than the average value for the population then parents chosen from this sample will give a larger-than-expected response to selection. The result this time is of an unexpectedly small response followed by an unexpectedly big one - again giving the negative correlation actually obtained. This explanation of the observed results is of course pure speculation until it can be shown by theory or by experiment that breeding values of a population do indeed vary in the manner suggested, namely being least in the direction of selection. Perhaps the well-accepted fact that heritability decreases from generation to generation as selection is continued in a given direction, is sufficient implication that heritability would decrease in the same direction within any one generation, and for the same reasons.

Whether or not the explanation given above is correct the fact remains that a negative correlation has been observed between the means of adjacent

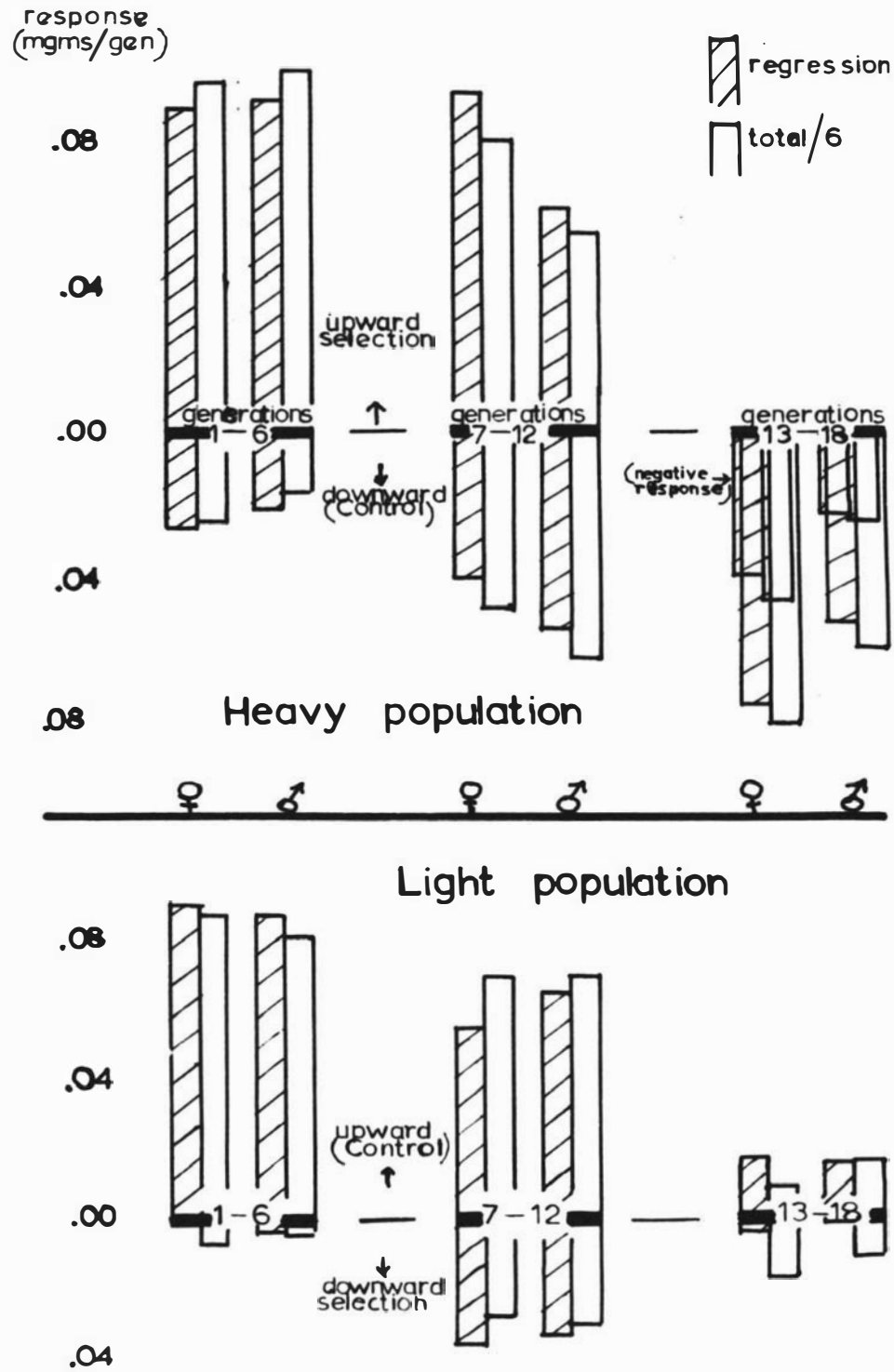
generations (though it was far less marked in a similar analysis of the Light population) and that because of this the measurement of average response by linear regression is strictly speaking invalid. The alternative method of measuring selection response is to divide the total response over the period by the number of generations in the period. Under this system the population mean weights of only Generations 0 and 6 are used and this of course makes the estimate free of any bias caused by correlation between successive generations. It may however rightly be claimed that this system conversely makes no use of valuable information provided by these intermediate generations.

The discussion above outlines the main theoretical difference between the two methods of measuring average selection response but it would be valuable to learn how greatly this difference is manifested in the actual data. For this reason the mean response per generation has been measured by both methods and the results are given in Table 9. The difference in value between each member of a pair of results is given, and also expressed in terms of a percentage of the second method estimate and this percentage is also contained in the table. However, such percentages are misleading unless considered in relation to the means themselves and this is most easily done by presenting the same data diagrammatically as well (Figure 4). Finally it should be recalled that for both methods the mean response has been calculated for each replication separately (sexes pooled) and then the results from all six replications have been added together and averaged. If the original data from all six replications had been pooled <u>before</u> regression was applied the result would have been different from that given whenever there was a gap in the data. Since such a gap was indeed present for one generation when Replication 5 died in the Light population this pooling did not take place.

Table 9: Two methods of measuring mean selection response

Generations	1 to 6		7 to 12		13 to 18	
	Mean response (mgms/generation)					
	a) Heavy population					
UPWARD SELECTION	♀	♂	♀	♂	♀	♂
regression method	.0898	.0920	.0936	.0622	-.0394	-.0234
total response/6	.0974	.1011	.0812	.0555	-.0475	-.0252
difference	-.0076	-.0091	.0124	.0067	.0081	.0018
difference/2nd estimate	-8%	-9%	15%	12%	17%	7%
	average difference = + .0020/.0437 = 5%					
DOWNWARD (CONTROL)						
regression method	.0269	.0218	.0409	.0548	.0762	.0523
total response/6	.0251	.0168	.0485	.0630	.0815	.0602
difference	.0018	.0050	-.0076	-.0082	-.0053	-.0079
difference/2nd estimate	7%	30%	-16%	-13%	-7%	-13%
	average difference = -.0037/.0492 = -8%					
	b) Light population					
UPWARD (CONTROL)	♀	♂	♀	♂	♀	♂
regression method	.0883	.0854	.0546	.0634	.0177	.0160
total response/6	.0849	.0786	.0681	.0681	.0095	.0162
difference	.0034	.0068	-.0135	-.0047	.0082	-.0002
difference/2nd estimate	4%	9%	-20%	-7%	86%	-1%
	average difference = 0/.0542 = 0%					
DOWNWARD SELECTION						
regression method	.0014	.0037	.0365	.0353	.0038	.0015
total response/6	.0075	.0045	.0283	.0301	.0175	.0111
difference	-.0061	-.0008	.0082	.0052	-.0137	-.0096
difference/2nd estimate	-81%	-18%	29%	17%	-78%	-86%
	average difference = -.0028/.0165 = -17%					

Fig 4: METHODS OF MEASURING AVERAGE RESPONSE



On most occasions the difference between the two estimates is less than 30%, although it is of course mostly personal as to whether 30% is considered to be a big or a small difference. The fact that the percentage differences are mostly quite similar to one another is due more to parallel changes in numerator and denominator than to constancy in either of them. On the few occasions that the difference between a pair of estimates is almost a hundred percent the cause is invariably a very low denominator rather than a very high numerator.

Finally, there is no obvious pattern in the direction or size of differences. They do not change systematically either in absolute value or percentage value from one period to the next, nor does either system of calculating mean response consistently give a higher value than the other system. However, this was not considered to be surprising as there were no obvious genetical or statistical reasons to expect such patterns.

Fluctuations of response: The fluctuation of a response about its mean value is a very important item of analysis whenever selection is continued for more than one generation. Indeed the comparison of two means or even the interpretation of one mean has very little statistical value unless this fluctuation is taken into account. There are again two main ways to measure this fluctuation between generations; either it can be calculated from deviations about the regression line and expressed as square root of error variance, or it can be calculated simply from deviations about the mean and expressed as standard deviation of response. The main theoretical difference between these two systems is this time apparently concerned with errors of weighing as well as sampling error. If the response was completely linear then there would be no fluctuation either about the mean response or about the regression line.

If on the other hand fluctuations do arise by mutation, genetic drift, natural selection or atypical sampling, that is genetic fluctuations which are perpetuated in the replications, they will be measured unequally by the two systems. But fluctuations which are not so fixed thereafter in the replication will receive even more unequal emphasis by the two methods. For example fluctuations caused by weighing error or by environmental changes are involved only once in the calculations by regression analysis. This is evident from the formula used in the calculation, namely

$$\sqrt{\text{error variance}} = \sqrt{\left[\frac{1}{n-2}\left\{\Sigma(y_i-\bar{y})^2 - \frac{[\Sigma(x_i-\bar{x})(y_i-\bar{y})]^2}{\Sigma(x_i-\bar{x})^2}\right\}\right]}$$

where n = number of generations

y_i = mean of Generation i

x_i = number of Generation i

Each generation mean (y) is used only once for calculating sums of squares and sums of products. However, the calculation of standard deviation of response uses each generation mean twice, thus

R_1 = first response = $y_1 - y_0$

R_2 = second response = $y_2 - y_1$ and so on.

Standard deviation of response = $\sqrt{\left[\frac{1}{N-1}\Sigma(R-\bar{R})^2\right]}$

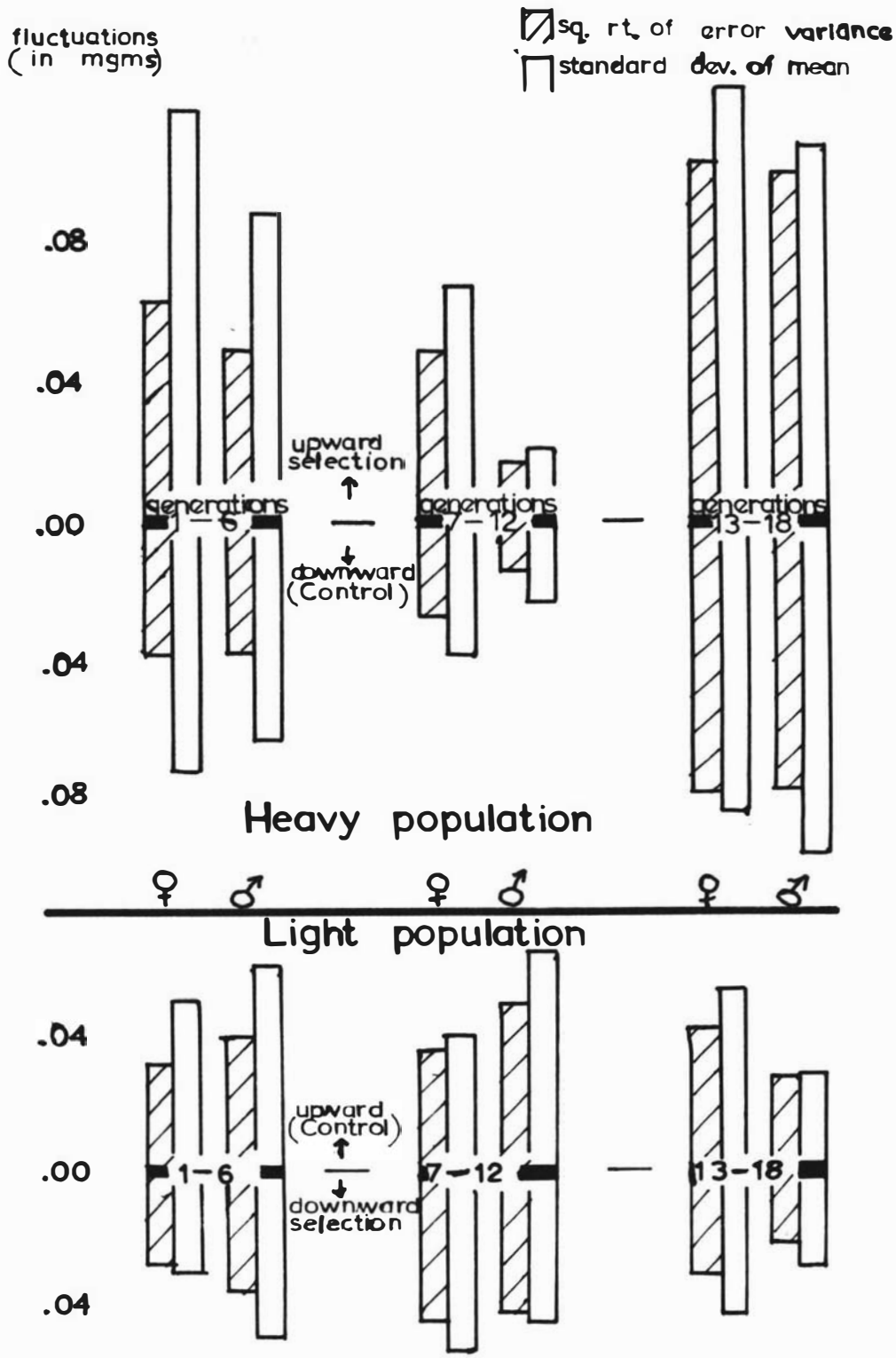
where N = number of generation responses = n - 1

Hence any errors of weighing involved in estimating y_1, y_2, etc have twice the effect on standard deviation that they have on root error variance. This would therefore tend to give the former a larger value than the latter, and the extent to which this actually happens is shown in Table 10 and Figure 5. It is pointed out that to simplify the computations for this table the data from all six replications were pooled in each generation before either the regression or response variances were calculated.

Table 10: Two methods of measuring response fluctuations

Generations	1 to 6		7 to 12		13 to 18	
	Fluctuations (mgms)					
	a) Heavy population					
UPWARD SELECTION	♀	♂	♀	♂	♀	♂
sq.rt. of error variance	.0658	.0493	.0482	.0167	.1050	.1022
standard deviation	.1204	.0894	.0689	.0215	.1266	.1101
difference	-.0546	-.0401	-.0207	-.0048	-.0216	-.0079
difference/2nd estimate	-45%	-45%	-30%	-22%	-17%	-7%
	average difference = -.0249/.0895 = -28%					
DOWNWARD (CONTROL)						
sq.rt. of error variance	.0385	.0380	.0278	.0158	.0789	.0774
standard deviation	.0727	.0639	.0391	.0218	.0857	.0970
difference	-.0342	-.0259	-.0113	-.0060	-.0068	-.0196
difference/2nd estimate	-47%	-41%	-29%	-28%	-8%	-20%
	average difference = -.0173/.0634 = -27%					
	b) Light population					
UPWARD (CONTROL)	♀	♂	♀	♂	♀	♂
sq.rt. of error variance	.0318	.0393	.0351	.0491	.0418	.0279
standard deviation	.0507	.0607	.0401	.0640	.0538	.0281
difference	-.0189	-.0214	-.0050	-.0149	-.0120	-.0002
difference/2nd estimate	-37%	-35%	-12%	-23%	-22%	-1%
	average difference = -.0121/.0496 = -24%					
DOWNWARD SELECTION						
sq.rt. of error variance	.0282	.0362	.0458	.0417	.0309	.0226
standard deviation	.0309	.0498	.0543	.0451	.0431	.0287
difference	-.0027	-.0136	-.0085	-.0034	-.0122	-.0061
difference/2nd estimate	-9%	-27%	-16%	-8%	-28%	-21%
	average difference = -.0077/.0420 = -18%					

Fig 5: METHODS OF MEASURING RESPONSE FLUCTUATIONS



For this parameter of variability the estimates from regression analysis are invariably smaller than those obtained by the other method and the difference ranges from 1% to 47% of the second estimates. There is a tendency especially in the Heavy population for this percentage difference to decrease from the first period to the third period but the significance of this trend is not known. Nor is there an obvious genetic reason for such a pattern to occur, for neither the denominators nor the numerators show a pattern of any sort. Nevertheless and regardless of any such patterns the degree of difference between any pair of estimates is mostly big enough to cause concern.

The cure for such concern is to standardize the thesis by using only one system of measuring variability throughout and the choice of which system to adopt will be made after discussing the measurement of response trends in the next section.

Trends of response: The two previous sections have discussed different methods of measuring the mean response and variability about this mean. It would be very useful to learn how much of this variability between generations results from trends in the sense that the response is gradually speeding up or falling off, and how much is due to fluctuations caused by genetic drift, weighing errors and environmental changes. It is assumed that any such trends would be caused by genetic changes induced by artificial and natural selection, and not by parallel trends of improvement or deterioration in the environment. There was no evidence for such environmental trends in the random mating type of Control in Experiment Two. Regardless of this however, it was considered that since the fluctuations obtained in the previous section were so big compared with their means, they would probably mask such trends in terms of statistical significance. It would be far better to reduce these fluctuations before attempting to determine the significance of any trends. Accordingly it

was decided not to measure these trends without first pooling the data of a selection line and its Control in each replication. The effect of incorporating this Control is to cancel all those factors which cause approximately equal but opposite variation in selection line and Control, such factors as environmental changes, scale, and natural selection. The fluctuation then remaining is caused by artificial selection, sampling and weighing errors, and any other forces which do not affect selection line and Control equally and oppositely. This remaining fluctuation may be small enough to allow some significant trends to emerge.

There does not appear to be any great choice of method in measuring these trends. Of the two methods given above for measuring means and fluctuations only the one method which involves generation responses can provide a measure of any trends. The alternative method which involved the fitting of a regression line to generation means apparently cannot do so, nor can the fitting of a regression line to accumulated divergence of response. But the first method, in which the response of each generation is separated from the other responses and in which the generation means themselves are not directly used, provides such a measure merely by fitting a regression line through the six separate responses of each replication. If the resulting regression coefficient is positive in sign then this indicates that the selection response has increased during the six generations, and if it is negative then the response is apparently falling off. The significance of these trends will be measured by the significance of the regression coefficient itself.

This now completes the discussion on the three main parameters of selection response; mean, fluctuations, and trends. It also completes the discussion on which methods are available to measure these parameters, and three main points

have emerged. Firstly, the two methods tested often gave fairly different results; secondly the regression method using generation means was unable to measure trends of response; thirdly this same method did not obey the assumption of non-correlation between adjacent generation means. It did obey the assumption that non-linearity would be insignificant, at least in the guinea-pig sample tested, but from a quick survey this non-linearity was much more noticeable in the response curves of Experiment Two. The one obvious advantage of this method was that in a least squares sense it was still the better method of measuring the average rate of response, regardless of the non-correlation premise. It is for this reason that the results of both methods have been presented in detail so far and both methods will again be considered in the prediction section that follows later in the thesis. However, for the sake of compactness the remaining results of the present section will involve only the non-regression method - which uses as units of analysis the generation responses, each obtained by subtracting one generation mean from the succeeding generation mean. For each replication these six responses in a period are averaged to give the mean response, their standard deviation about this mean gives a measure of the fluctuations, and the regression line fitted to the six of them gives a measure of the trends. Finally it is pointed out that in all cases the six replications are treated independently, so that a mean, standard deviation, and regression coefficient is calculated for each one; but within each replication the female and male data are pooled before such calculations are made. The overall value of mean, standard deviation or regression coefficient for any selection period is then obtained by simply averaging the six values from the six replications. The variation between replications in these values will be discussed either at the time or later.

(c) THE SELECTION RESPONSE

Response mean: The mean responses for each period are given in Table 11. Each mean value given has been obtained by averaging the six replications but each replication is itself the mean of six generations. Because of this latter fact the six replication means can be considered to lie on a normal distribution curve. The numbers of points involved (six) is too low to give an exact proof of this normality but the Central Limit Theorem justifies such assumption being made. By virtue of this assumption the standard error of each period mean has been calculated from the variability of replication means about it, and these standard errors are given in the table. By making use of these standard errors the differences between means of adjacent periods and between those of main lines and Controls have been tested for statistical significance. This significance is calculated from the t-test formula

$$d = \frac{\mu_a - \mu_b}{\sqrt{\frac{s_a^2 + s_b^2}{n}}}$$

where μ = period mean

s^2 = variance among replication means

n = number of generations in each period.

The significance of d is measured from Student's "t tables" and the degrees of freedom are determined from Bailey (1959) thus

$$\text{d.f.} = \frac{1}{\frac{u^2}{n_1 - 1} + \frac{(1-u)^2}{n_2 - 1}}$$

where $u = \dfrac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

in which the two variances are not assumed to be equal. This significance is expressed as a probability that the difference is due to chance and the value of this probability is given in brackets between each pair of means.

Table 11: Mean response in each period (mgms per generation)
( ) = Statistical probability that the means on either side are not significantly different

Generations	1 to 6	7 to 12	13 to 18
		Heavy population	
upward selection	.0993 ± .0299(.5)	.0685 ± .0207(.02)	-.0364 ± .0269
	(.1)	(.7)	(.02)
downward (Control)	.0211 ± .0220(.3)	.0558 ± .0161(.6)	.0711 ± .0200
		Light population	
upward (Control)	.0817 ± .0203(.7)	.0693 ± .0209(.05)	.0129 ± .0116
	(.02)	(.2)	(.9)
downward selection	.0063 ± .0157(.4)	.0287 ± .0152(.5)	.0142 ± .0112

In the Heavy population the upward response decreases with time and the downward (Control) response increases. This pattern could be explained by postulating that natural selection operates against heavier pupae, in which case it would oppose artificial selection in the upward line and assist it in the downward Control, these effects of natural selection increasing through the second and third periods of selection as the population mean weight itself rises. However, the explanation is unlikely to be as simple as this for two reasons; firstly the pattern would be much less distinct without the very large fall in mean weight at Generation 18 (and this large sudden fall is more likely due to a change of environment rather than to natural selection). Secondly the

upward line gives a much greater response than its Control during the first six generations so this asymmetry requires a separate hypothesis. The original hypothesis of natural selection requires amendment, namely that such natural selection apparently begins to operate only at a very high pupa weight, well above that of the original base population.

In the Light population the upward (Control) response again decreases significantly with time. There is an increased response downwards but it is neither regular nor significant. This overall pattern once more could be explained partly by the action of natural selection against heavy pupae if it is also accepted that the overall selection response is gradually being curbed by the loss of genetic variation. This explanation however, still leaves the problem of asymmetrical response, for as with the Heavy population there is considerably greater progress in the upward direction than downwards during the first two periods. The asymmetry phenomenon is treated in detail by Falconer (1961) and Clarke et al., (1961) and they suggest several theoretical causes of it. In the present experiment the consistently greater response upwards suggests that scale effects may be partly involved and so efforts were made accordingly to remove this effect from the data.

Falconer (1961) explains that any asymmetry of the response that is caused by exponential growth scale effects can be removed by converting the data to logarithms. This would be an extremely laborious task if each measurement had to be converted separately but Falconer supplies a conversion formula that uses only generation means and yet is still quite accurate; this formula is

$$\overline{\log x} = \log \bar{x} - \tfrac{1}{2} \log (1 + C^2)$$

where $\bar{x}$ = population mean weight

C = coefficient of variation.

Table 12: Asymmetry of response

Generations	1 to 6		7 to 12		13 to 18	
	♀	♂	♀	♂	♀	♂
A) mean weights (mgms)						
Heavy population						
upward selection	2.63	2.51	3.10	2.91	3.28	3.11
before selection began	2.25	2.13	2.84	2.73	3.32	3.07
downward (Control)	2.19	2.08	2.67	2.50	3.13	2.95
Light population						
upward (Control)	1.95	1.81	1.79	1.67	1.50	1.40
before selection began	1.62	1.50	1.57	1.47	1.47	1.34
downward selection	1.57	1.48	1.44	1.33	1.37	1.27
B) coefficients of variation (%)						
Heavy population						
upward selection	8.7	9.3	7.9	9.1	7.0	7.7
before selection began	8.7	9.6	8.5	8.9	6.3	8.7
downward (Control)	10.1	10.3	9.5	10.6	7.2	8.4
Light population						
upward (Control)	10.5	10.8	11.0	11.0	12.6	11.6
before selection began	13.3	13.0	12.3	12.6	12.6	12.3
downward selection	12.1	12.7	12.4	12.6	13.1	13.0
C) mean of logarithmic weight ($\log_{10}$)						
Heavy population						
upward selection	.419	.398	.490	.462	.515	.492
before selection began	.351	.326	.452	.435	.520	.485
downward (Control)	.338	.316	.425	.396	.494	.468
Light population						
upward (Control)	.288	.255	.250	.220	.173	.143
before selection began	.206	.172	.193	.164	.164	.124
downward selection	.193	.167	.155	.120	.133	.100

The degree to which this conversion formula reduces the asymmetry of response is illustrated in Table 12 above. The table is divided into three sections, the first one containing the raw data of mean weights before conversion, the second one containing the appropriate values of coefficient of variation, and the third section containing the appropriate data after conversion. Each of these three sections itself contains three rows of data, the middle row referring to the population value for the generation beginning a period of selection, the upper row referring to the population value averaged over the subsequent six generations of upward selection, and the lower row referring to the population value averaged over the subsequent six generations of downward selection. On every occasion the values given in the first two sections of the table are obtained by averaging the mean weights and coefficients of variation from all replications and it is these average values given in the table which are substituted in Falconer's formula.

The feature to look for in Table 12 is asymmetry of this converted logarithmic data in section (C); that is, does the difference between starting value and upward value exceed that between starting and downward value? Study of the table shows that such asymmetry is indeed present in both populations in Period 1 and in all but the Heavy population male data of Period 2. In Period 3 the asymmetry is in the opposite direction. If these results are now compared with those of Section (A), which involved the unconverted mean weight data it can be seen that the patterns of asymmetry are exactly the same in both tables: where the upward response was greater in one table it continues to be greater in the other. This result means that transformation of the data to logarithms has qualitatively achieved nothing; there is no removal of the scale effect and so it cannot be claimed that scale effect is an important cause of the asymmetry. In a quantitative sense there may have been some degree of reduction in this asymmetry but it is very difficult to decide on

what basis such reduction could be measured. The only obvious system would be to measure the asymmetry in each case in terms of a ratio, upward response to downward response, and then to observe the change of ratio when the data was transformed to logarithms. However, it was not clear to the author that this system of measuring the change in asymmetry was indeed meaningful and so it was not used. Finally, perhaps if the coefficients of variation had been considerably larger this transformation to logarithms would have had more effect (Falconer, 1961).

The second method mentioned of cancelling a scale effect is by converting the response data to the parameter known as realised heritability. This parameter is calculated as a fraction and its merit in the present context relies on the assumption that any scale effect which may distort the numerator (selection response) of this fraction may also distort the denominator (selection differential), and by a proportional amount in the same direction.

Table 13 contains the realised heritability data for each period and once again the value given for a period is the mean value of 72 separate heritabilities (2 sexes X 6 generations X 6 replications). Each replication mean heritability is therefore obtained by averaging 12 items and the Central Limit Theorem is used to assume that the 6 replication means would therefore lie on a normal distribution curve. This assumption allows the author to calculate standard errors based on the variability among replication means, and these standard errors are presented in the table. The statistical probability that differences between periods or between selection lines and their Controls are not significant is recorded in brackets between each pair of adjacent means. The probability figure may be used to ass ss if the asymmetry greatly reduces when the scale effect is removed, for in the event that it does so this probability figure would surely be larger in value than the corresponding figure calculated for the response data in Table 11.

Table 13: Realised heritability in each period

Generations	1 to 6	7 to 12	13 to 18
		Heavy population	
upward selection	.282 ± .092(.6) (.1)	.213 ± .059(.02) (.4)	-.131 ± .088 (.02)
downward (Control)	.060 ± .072(.4)	.135 ± .042(.4)	.197 ± .055
		Light population	
upward (Control)	.258 ± .073(.7) (.05)	.209 ± .059(.1) (.2)	.045 ± .046 (.9)
downward selection	.016 ± .053(.4)	.088 ± .060(.7)	.047 ± .052

The heritability patterns that arise in the table above are exactly the same as those obtained for mean response in Table 11, and the probability of significant differences between periods is also very much the same. These probability figures between a selection line and its Control do not show the expected increase of value and indeed two of the differences actually show the opposite trend. It must therefore be concluded that these heritability data suggest that a marked logarithmic scale effect is not present. It would not be disputed that some scale effect is present in the experiment, for the graphs of phenotypic variance strongly suggest this, but the analysis has been unable to prove that the logarithmic cause is significant in amount. This judgment means that some other explanation must be sought for the strong asymmetry in both Heavy and Light populations during the first period of selection. The same type of asymmetry still persists in the second period of selection but it is not so marked, and it has disappeared or reversed in the third period. While it is true that differences between upward and downward period means

seldom exceed the standard significance level of .05 this is regarded as being due to inadequate replication rather than proof that such patterns of asymmetry do indeed not exist. The attitude adopted by the present author to the asymmetry is that it is indeed a real and not a chance result, and that it may be due to non-logarithmic scale effects. But it must be admitted that in the absence of decisive tests such suggestions are merely speculation and no simple hypothesis appears able to cater for both populations and all three periods of selection. Natural selection may be involved to some extent but there is little tendency in Table 2 for generation interval to increase in the Light population or the Heavy population. Also the fact that conversion of the data to realised heritability has not removed the asymmetry, alongside the argument that it should remove most of any such asymmetry caused by natural selection, means that natural selection is not a large factor in the asymmetry. Falconer (pages 212-5) discusses other possible causes of asymmetry such as "directional dominance" but in the absence of concrete tests this discussion will not be repeated here. The question of what is definitely causing this asymmetry must be left open.

The final point of interest about asymmetry encountered in the results of the present experiment is its steady decrease from period to period. The extent to which this decrease merely reflects that the genetic variation is itself falling can be measured roughly by pooling the data of a selection line with that of its Control. In this way the fluctuations due to environment and the trends of asymmetry are mostly cancelled out, leaving behind a much more reliable estimate of the true genetic response. Since such response is theoretically related to the amount of genetic variation remaining in the population, a comparison of these responses in the three periods of selection will indicate if the genetic variation of the population is being reduced by selection. The first section of Table 14 contains the results of this process of calculation.

In each replication the total upward (or downward) response over six generations has been averaged with that of its Control. Then the six replications themselves have been averaged to give a mean response, and their variability about this mean is presented as a standard error. The same data is presented diagrammatically in Figure 6 to show the replication differences.

Table 14: Response with Control incorporated (mgms per generation)

Generations	1 to 6	7 to 12	13 to 18
		Heavy population	
a) mean response	.0601 ± .0071 (.9)	.0620 ± .0038 (.01)	.0172 ± .0119
		Light population	
	.0438 ± .0073 (.8)	.0494 ± .0166 (.1)	.0136 ± .0033
		Heavy population	
b) standard deviation	.0635	.0475	.0550
		Light population	
	.0732	.0565	.0409
		Heavy population	
c) regression coefficient	-.00253 (.7)	+.00162 (.3)	-.00621
		Light population	
	-.00800 (.7)	-.00377 (.8)	-.00636

The mean response shows almost no change from Period 1 to Period 2 but then it drops drastically in both populations. This pattern is contrary to that expected from additive genetic theory. If the selection response had been due entirely to the increasing frequency of genes with the highest additive

Fig 6: REPLICATION DIFFERENCES
(a) PERIOD PARAMETERS

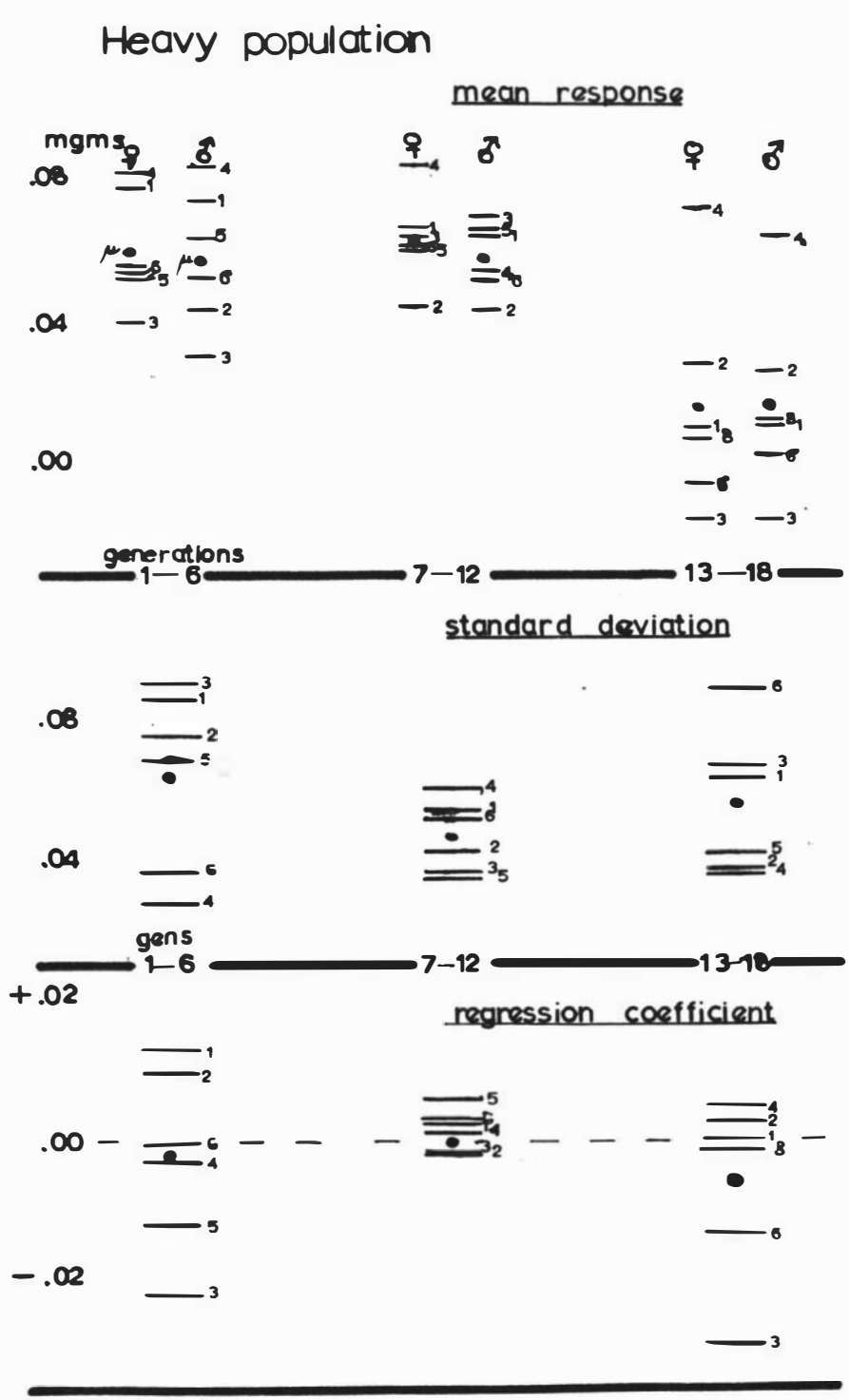
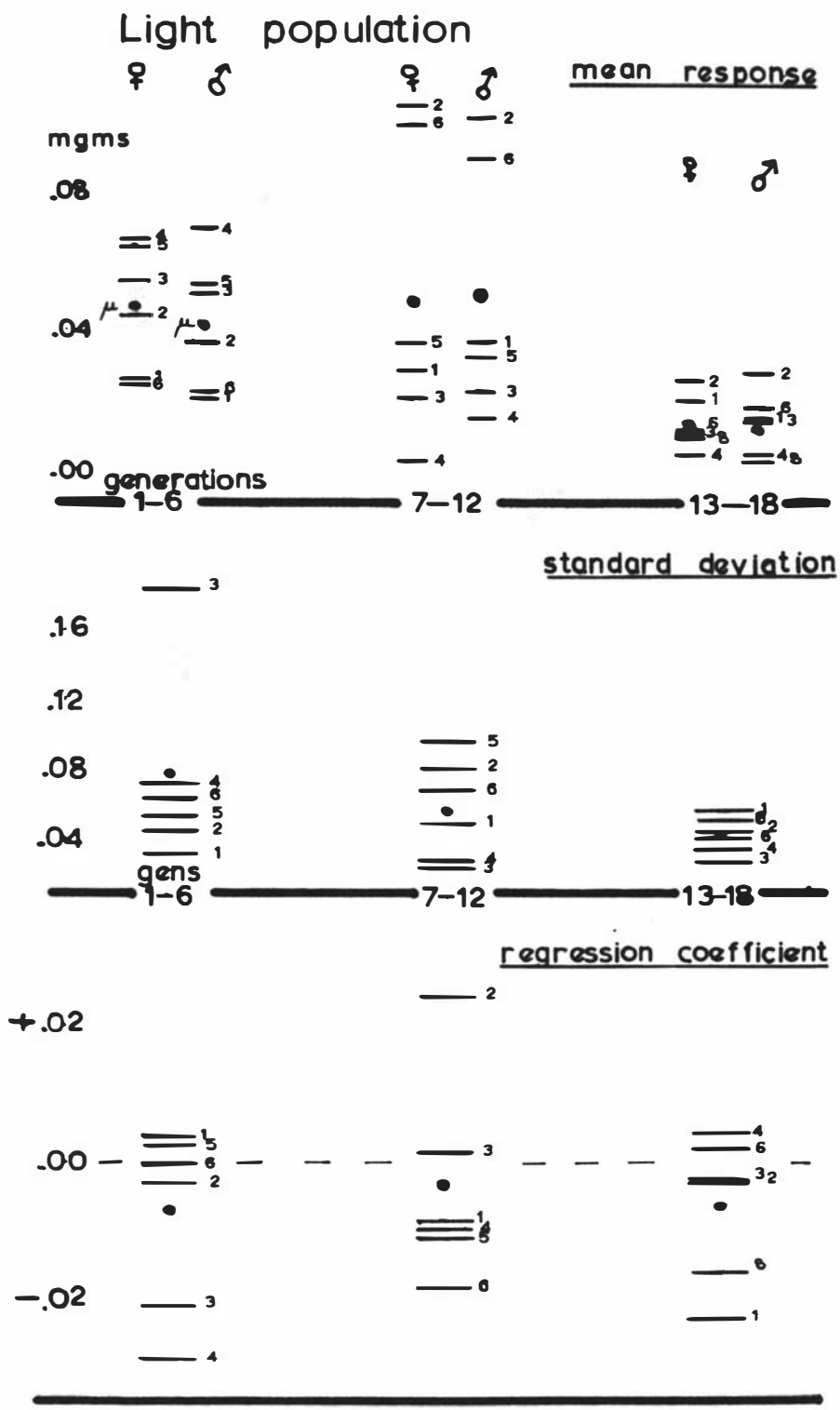


Fig 6 : continued



values, and if such genes had been at intermediate frequencies when selection began, then their rate of increase would be greatest in the first few generations. This rate of increase would gradually fall off in later generations. Since the population value theoretically depends on the frequency of these genes the mean response should likewise be greatest in the first period of selection and gradually tail off in later periods. The fact that it does not appear to do so suggests that gene action is not entirely additive. There would be very little doubt that the problem of interpreting this pattern of response is related to the problem of explaining the asymmetry encountered earlier, because exactly the same data is involved in each case and almost exactly the same treatment of this data. Like the asymmetry question it must be left fairly open at this stage.

Standard deviations: The estimate of standard deviation for each period is given in Figure 6 and section B of Table 14. The method used to obtain this estimate has been discussed earlier but will be explained again. In each replication the female and male data were pooled. Then the responses of a selection line and of its Control were averaged for each generation in turn, to give an estimate of the genetic response for that generation. Next the generation responses themselves were averaged and their variability about this average was measured in terms of standard deviation. The final step was to average the six standard deviations, one from each replication, and it is this value which appears in the table.

The resulting standard deviations are mostly much bigger than their corresponding mean responses higher in the table, and this illustrates how big are the fluctuations in an individual replication. Above all it stresses that the relatively smooth responses obtained by pooling all replications in Figure 1

depend for their smoothness on high replication. There are probably several factors operating to cause such fluctuations and these factors obviously must have unequal effects on a selection line and its Control, otherwise the effects would be cancelled out. On this basis the factors of natural selection, scale, and environmental change would probably contribute very little to the values of standard deviation above, and the main contribution to these values probably comes from sampling and weighing error, genetic drift, and asymmetrical effects of selection.

It is not clear if the value of standard deviation should be expected to increase or decrease systematically from the first period through to the third. On the one hand the fact of a decreasing genetic variability might give reason to expect smaller fluctuations from genetic drift and sampling error: on the other hand it is reasonable to expect that the inbreeding level rises as this genetic variability decreases. A highly inbred population is popularly considered to be poorly buffered against environmental change and such effect would be manifested in big fluctuations of response. The net effect of these two confounding factors is hard to predict, and perhaps a systematic pattern of increase or decrease should not be expected. As it turns out no such pattern is obtained in the Heavy population anyway. In the Light population however, the value of standard deviation decreases from one period to the next and this trend ought to be analysed for significance. The first step in such a test of significance is to ascertain that within any one period the six standard deviations (one from each replication) are reasonably homogeneous. If they are not so, then there is little meaning to the mean value of these six standard deviations and there is even less point to comparing the mean value of one period with the mean of the next period. However, if the six standard deviations of any one period are similar in value they may be averaged to give an adequate description of variability for that period. The test actually applied to each

period was Bartlett's test for the homogeneity of several variances, which required the standard deviation data to be converted back to variances. Then the following expression was calculated from these six variances.

$$\frac{1}{C}\left\{ f \log_{10} s^2 - \Sigma f_i \log_{10} s_i^2 \right\}$$

where $C = 0.4343\left[1 + \frac{1}{3(k-1)}\left\{\Sigma\frac{1}{f_i} - \frac{1}{f}\right\}\right]$

k = number of replications = 6

n_i = number of observations in ith rep = 6

$f_i = n_i - 1$

$f = \Sigma f_i$

$s^2 = \Sigma f_i s_i^2 / f$

s_i^2 = variance of ith replication.

The top expression is distributed approximately like χ^2 with (k-1) degrees of freedom. Bailey (1959) claims that the main fault of this test is that it depends on the distributions in the six variances all being normal, as moderate departures from normality may cause the test to be misleading. In the present case each variance is computed from only six items, which is too few to provide a reliable test of normality. However, the fact that each item is itself the mean of 64 responses fairly assures this normality (Central Limit Theorem).

The results of Bartlett's test for homogeneity are given below in Table 15.

In the Light population two of the three expressions have significant P-values, which indicates that the six variances are heterogeneous. It is therefore fairly meaningless to pool them in order to compare the average variances <u>between</u> periods and so it must be concluded that there is no satisfactory

Table 15: Test for homogeneity of variances

(E = value of $\frac{1}{C}\left\{ f \log_{10} s^2 - \Sigma f_i \log_{10} s_i^2 \right\}$

P = probability that E is not significant)

Generations	1 to 6		7 to 12		13 to 18	
	Heavy population					
	E	P	E	P	E	P
	8.45	.1	2.87	.8	6.22	.3
	Light population					
	20.57	.01	12.93	.05	3.16	.7

test for deciding if the apparent trend for standard deviation to decrease in value from Period 1 through to Period 3 is really significant.

In the Heavy population there is much more homogeneity among the variances and it is therefore valid to pool them within each period. The comparison of pooled variances between periods is then made by an F-test, with the pooled degrees of freedom. These variances were (taken from s^2 in the expression above):

Generations 1 to 6 s_1^2 = .004586

7 to 12 s_2^2 = .002375

13 to 18 s_3^2 = .003373

The F-ratios are s_1^2/s_2^2 = 1.93 ns

s_2^2/s_3^2 = 1.42 ns

(The 5% significance value required is 2.07 on a 2-tailed test).

Therefore it must be concluded that although over both populations there are indications that the size of fluctuation as reflected by standard deviation decreases from one period to the next, this trend could not be proven by statistical analysis to be significant.

Regression coefficient: Suppose that in each replication the six responses averaged with Control (one per generation) of a period are plotted on a graph and a regression line is fitted to them. If the regression coefficient of this line is approximately zero it implies that there has been no significant change in the rate of response during selection. If the coefficient is positive or negative it implies that the response has respectively increased or decreased in later generations, and this in turn is a guide as to whether the genetic variability within a replication is increasing or decreasing very markedly.

The regression coefficient is shown separately for each replication in Figure 6 to allow comparison of replications with one another. In the event that they are homogeneous the six values can be pooled to give the average coefficient for a period and then these period means themselves can be compared. The analysis that follows (Table 16) is therefore a standard F-test of heterogeneity of regression. It was carried out for each period in turn in both populations.

The F-value does not approach significance on any occasion so it may be accepted that within any one period the regression coefficients from all six replications are fairly similar in value. The average value of regression coefficient for any one period may therefore be calculated from the pooled data and this value may now be compared between periods. The comparison is made

Table 16: Test for homogeneity of regression lines - Summary of F-tests

Generations		1 to 6		7 to 12		13 to 18	
		Heavy population					
Source of variation	df	MS	F	MS	F	MS	F
Differences between regressions	5	.00309	0.61	.00018	0.06	.00303	0.89
Deviations from indiv. regressions	24	.00506		.00292		.00340	
Deviations from av. regressions within groups	29	.00472		.00245		.00334	
		Light population					
Differences between regressions	5	.00335	0.38	.00371	1.00	.00147	0.83
Deviations from indiv. regressions	24	.00887		.00371		.00177	
Deviations from av. regressions within groups	29	.00792		.00371		.00172	

by t-test (at 29 d.f.) and the probability of a non-significant difference as indicated by this value of d is given in brackets between each pair of period means in section (C) of Table 14. These probabilities do not approach significance on any occasion and so it must be concluded that no significant differences have been found between periods in the size of their average regression coefficients.

Perhaps the most important result in this analysis is the observation from Figure 6 that most of the regression coefficients of individual replications are

negative in value. Each of these values is not significantly negative in itself but the reasonable consistency of negativity is firm evidence that within each period the genetic response gradually falls off from one generation to the next. But the analysis did not succeed in showing that such response falls off fastest during the first period of selection and slowest during the third period - as would be expected to be the case under the theory of additive gene action.

Finally it should be observed that the values of regression coefficients in Table 14 are very small compared with the corresponding standard deviations in the same table. This implies that any trends of response are only a very minor cause of the total fluctuation between generations.

(D) DIFFERENCES BETWEEN REPLICATIONS

The analysis of a selection response would be incomplete if no attempt was made to measure the importance of differences between replications. If no such measurement is made there could be very little meaning to the practice of averaging the responses of replications in a period, in order to compare one period mean with another period mean. In the earlier sections of the present thesis replication variability was taken into account when period means were being compared, and for the parameters of standard deviation and regression coefficient the replications were shown to be sometimes heterogeneous, sometimes not. Because of this heterogeneity any comparison between different periods could not be taken very far by standard statistical methods.

An alternative method of describing the variability between replications in any one period, and of comparing this variability between two different periods, is presented below in Table 17. It can be applied to each of the three parameters, mean response, standard deviation, and regression coefficient; but the results must always be interpreted with caution. The method consists simply of calculating the deviation of each replication from the period mean value of all six: and from this, the standard deviation for each of the three parameters in turn. It might perhaps be confusing to envisage a standard deviation of six standard deviations but it is considered that since all three parameters are first-order derivatives of the data, then the Central Limit Theorem gives meaning to the standard deviations of all three. Table 17 also contains the values of coefficient of variation, calculated from these standard deviations and their respective period means - again for all three parameters in turn.

The coefficient of variation between replications has a fairly large value for all three parameters but it is particularly large for the case of regression coefficient. A survey of the table shows that this latter case is not because

Table 17: Replication differences 1) period parameters

Generations	1 to 6			7 to 12			13 to 18		
	Mean	S.D.	C.V.	Mean	S.D.	C.V.	Mean	S.D.	C.V.
a) mean response				Heavy population					
	.0601	.0175	29%	.0620	.0093	15%	.0172	.0292	170%
				Light population					
	.0438	.0184	42%	.0494	.0408	83%	.0136	.0082	60%
b) standard deviation				Heavy population					
	.0635	.0257	40%	.0475	.0118	25%	.0550	.0205	37%
				Light population					
	.0732	.0549	75%	.0565	.0298	53%	.0409	.0106	26%
c) regression coefficient				Heavy population					
	.00253	.01331	526%	.00162	.00322	199%	.00621	.01301	210%
				Light population					
	.00800	.01383	173%	.00377	.01489	395%	.00636	.01098	173%

of large deviations in an absolute sense, but merely because the period means themselves are always near zero.

There are no obvious patterns of changing variability from one period to the next, either in the deviations themselves or in the coefficients of variation.

There are other aspects in which the variation between replications may change systematically as selection proceeds, and perhaps the best way to detect these patterns is to discard the Control data and analyse just the main selection lines from Generation 0 to Generation 18; this longer period allows more chance for such patterns to develop and be observed on a graph. In the analysis that follows the author is concerned with the mean weight of each replication, not with its response. The aim is to measure at each generation the variability

between replications in mean weight, and then to observe how this variability alters from Generation 0 through to Generation 18. Such analysis is quite rare in earlier literature and this is probably because adequate replication is quite rare also.

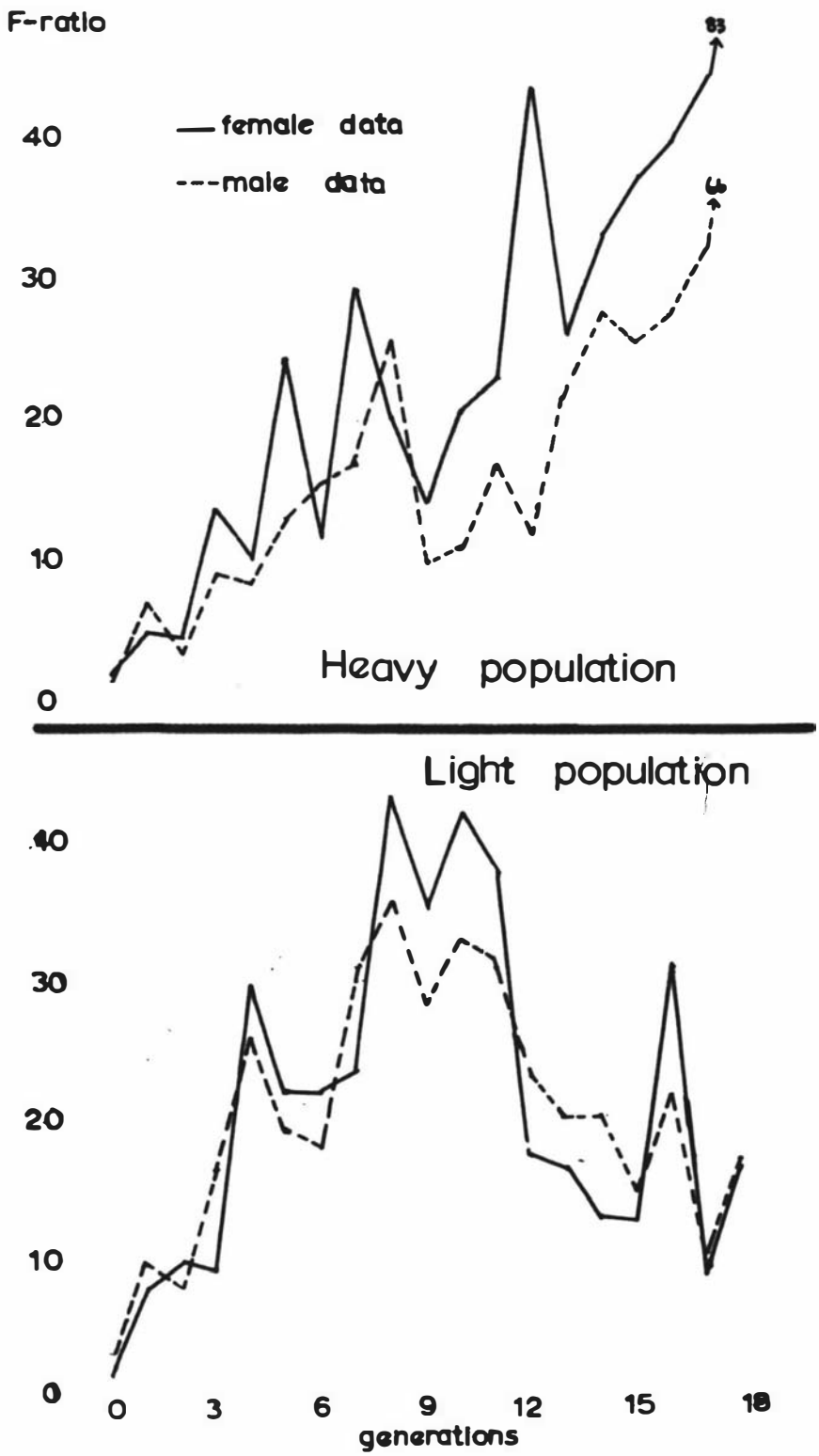
The terms in which this variability between replications is to be expressed requires some thought. If it is expressed as statistical variance no allowance is made for scale effects as the population mean rises or falls. On the other hand this scale effect was shown to be too small in earlier analyses to justify measuring the variability in terms of coefficient of variation between replications (Kojima and Kelleher, 1963). The author reasoned instead that this scale effect would be best removed by comparing the variance between replications to the variance within replications, that is as the F-ratio of a variance analysis. After all, this was the method used to show that there were no significant differences between replications when they were originally sampled from the base population; it should equally well indicate if such differences become more significant as selection proceeds. The form of variance analysis is set out below.

Source of variance	SS	df	MS	F	Variance Components
Between replications	A	5	A/5=C	$^{C}/_{D}$	$\sigma^2_W + 32\sigma^2_B$
Within replications	B	186	B/186=D		σ^2_W

The F-ratio ($^{C}/_{D}$) was calculated separately for each sex in every generation and the results are graphed in Figure 7.

One important point should be made about these graphs before they are analysed too far. This point is that in the Light population Replication 5 died out at Generation 12 and was thenceforth replaced by Replication 8. However, the former one had been easily the lightest replication of all six and so its removal from the variance analysis would automatically reduce the F-ratio,

Fig 7 : REPLICATION DIFFERENCES
(b) VARIANCE RATIO



perhaps substantially. The addition of Replication 8 to the variance analysis would not restore this F-ratio to its former value because this replication deviated less in weight from the population mean. This explanation probably accounts for the drop in F-ratio at about Generation 12 in the Light population.

Apart from the above oddity both populations record an erratic but obvious increase in F-ratio as selection proceeds; therefore the variability between replications is becoming more and more significant. The F-ratio has a 5% (1%) significance value of 2.3 (3.2) for 5 and 186 degrees of freedom and the fact that it exceeds this value from about the first generation is ample warning to use more than just one or two replications in selection experiments of this type. The differential effects of genetic drift and selection on the six replications are apparently important from the start and they apparently accumulate as the generations proceed.

Increasing variability between replications has two main causes. There would have been at least a few genetic differences present between them when they were originally sampled from the base population even though these differences were not proven to be significant. Firstly then, more differences would arise by the random action of genetic drift and so the replications would begin to diverge in genetic content. Secondly, the effects of artificial selection on these differences might perhaps cause them to diverge even more. It is possible to estimate the relative importance of selection and genetic drift in causing variability between replications but it must be stressed that such estimates are very rough, and based on many assumptions. The method used here is to isolate out of the variance analysis above, the between-replication component (δ^2_B). Then the <u>theoretical</u> value of this component under the action of genetic drift is calculated, and subtracted from the <u>observed</u> value.

The remaining variance which is not accounted for by genetic drift is thereby attributed to the differential effect of selection on the six replications.

The reason that such a system is not very precise relates to the formula used for estimating variance due to genetic drift. This formula is given by Falconer (1961).

$$\sigma^2_B = 2F\sigma^2_{add}$$

where F = inbreeding level of population

σ^2_{add} = additive variance in base population.

The accuracy of this formula depends on the accurate estimation of these two terms and experience from the literature and from later results in the present thesis suggests that it is unusual to obtain an accurate estimate of additive variance. The value used in the present case was obtained by averaging the sire and dam components of a diallel analysis on the base population, but its exact derivation will not be described until later. The following values were used for additive variance -

female data: σ^2_{add} = .00563

male data: σ^2_{add} = .00413

Estimation of the inbreeding level each generation may also be quite inaccurate, especially if there develops some natural selection against homozygosity. The inbreeding level is estimated from the rate of inbreeding and this rate in turn is calculated from a formula whose separate terms may be estimated with error. Nevertheless it was decided to continue this attempt to evaluate the variance resulting from genetic drift, but with firm appreciation of its fallibility.

Rate of inbreeding depends on the number of parents chosen in each genera-

tion and in the present experiment that number was eight. But only in an idealised population would these eight parents contribute with equal probability to the next generation, that is mate with one another in an entirely random manner and with equally random survival of their progeny; on many occasions some parents would contribute far more than others to the next generation, and this tendency increases the effective inbreeding rate. Such an effect can be measured if the variance of family size per parent is known but in the mass-mating system of the present experiment this could not be measured. However, since Shrode (1960) claims that there is no tendency for preferential mating in Tribolium the variance of family size may be quite small. One further source of error which had to be catered for was the tendency that in a selection character of high heritability those parents which would be selected had more genes in common than if, as the inbreeding theory assumes, they were a random sample of the population. Robertson (1961) has investigated this source of error and produced a formula to deal with it for the case of fullsib families.

$$N/N_e = 1 + 2\tau^2 h^2 \left[1 - h^2 \tau \left(\tau - x\right)\right]$$

where N = actual number of parents

N_e = effective number of parents

τ = selection intensity in standard measure

h^2 = heritability

x = point on asymptote of unit normal curve from which a perpendicular cuts off the proportion of animals selected to be parents.

In both Light and Heavy populations the proportion of animals selected was 4/32, which corresponds to

τ = 1.59 (Fisher and Yates, 1963, Table 20)

x = 1.2 (Fisher and Yates, 1963, Table 9(1))

The value taken for heritability was that which seemed to be about average over the eighteen generations, but disregarding the strongly negative response at Generation 18 in the Heavy population. These values were -

Heavy population: $h^2 = .15$

Light population: $h^2 = .05$

Substitution of these values in the formula gave -

Heavy population: $N/N_e = 1.69$

$\therefore N_e = 8/1.69 = 4.7$

Light population: $N/N_e = 1.25$

$\therefore N_e = 8/1.25 = 6.4$

(Note however, that since the mass-mating parents probably produce some halfsib families as well as fullsib ones, the true reduction in effective number of parents is probably not as drastic as shown above).

The respective rates of inbreeding ($1/2N_e$) used are therefore -

Heavy population: $\Delta F = .106$

Light population: $\Delta F = .078$

The inbreeding level at any generation t is then calculated as (Falconer, 1961) -

$$F_t = \Delta F + (1 - \Delta F)F_{t-1}$$

and the final step in estimating variance due to genetic drift is from the formula -

$$\sigma^2_B = 2F\sigma^2_{add}$$

Table 18 contains both of these parameters - the inbreeding level at each generation and the estimated variance due to drift, for each sex.

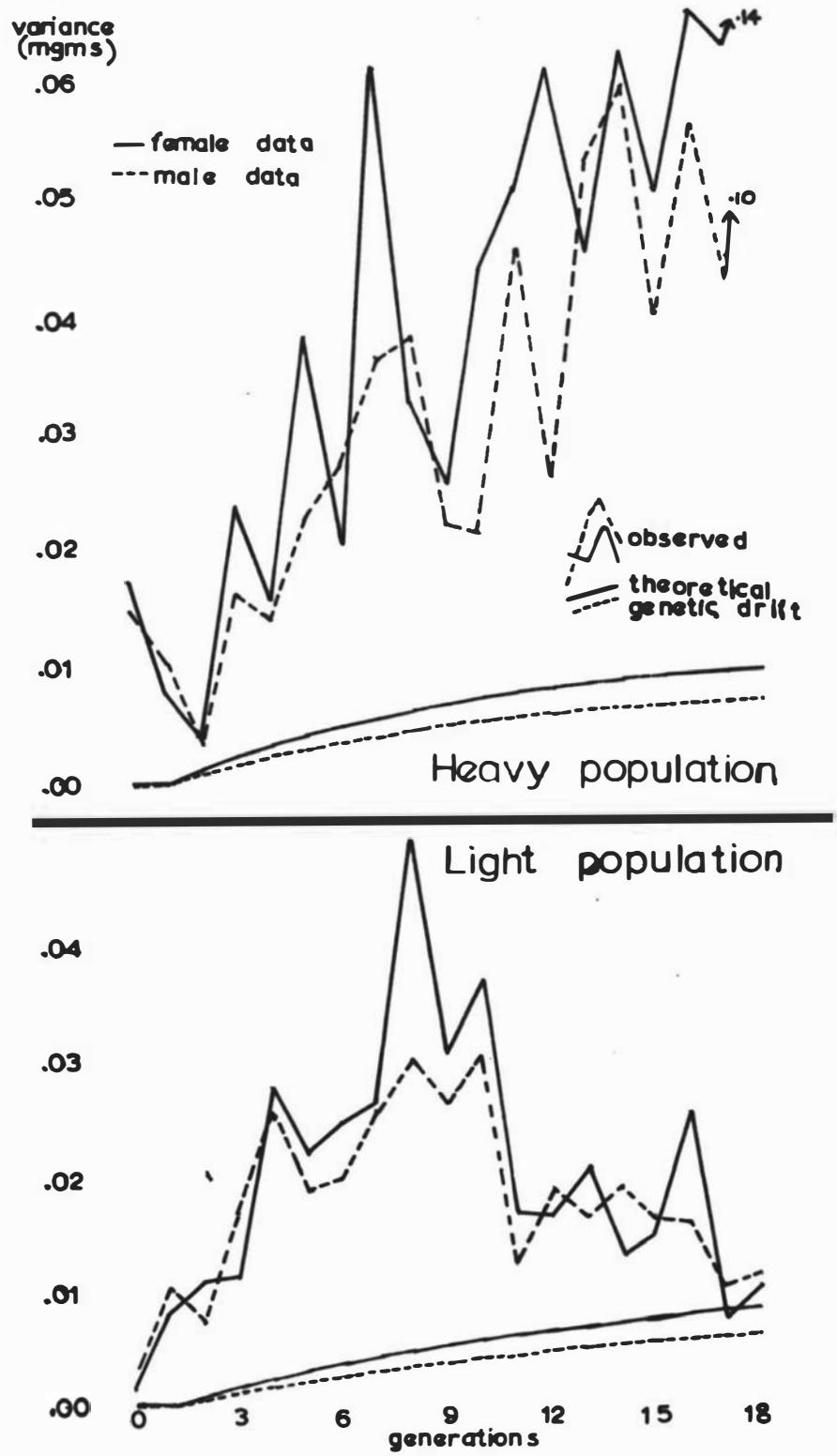
The variance data of this table are presented again as a graph in Figure 8. This enables them to be compared with the actual observed values of the between-replication variance component, which also appears in Figure 8. The conclusion

Table 18: Replication differences
ii) between-replication variance component

Generation	Heavy population			Light population		
	F	variance between reps		F	variance between reps	
		female data	male data		female data	male data
0	.000	.00000	.00000	.000	.00000	.00000
1	.000	.00000	.00000	.000	.00000	.00000
2	.106	.00120	.00088	.078	.00088	.00065
3	.201	.00227	.00167	.150	.00169	.00124
4	.286	.00323	.00237	.216	.00244	.00179
5	.362	.00409	.00300	.277	.00313	.00230
6	.430	.00486	.00357	.333	.00376	.00276
7	.490	.00554	.00407	.385	.00435	.00320
8	.544	.00614	.00452	.433	.00489	.00359
9	.592	.00669	.00491	.477	.00539	.00396
10	.635	.00718	.00527	.518	.00585	.00430
11	.674	.00762	.00559	.556	.00628	.00461
12	.709	.00801	.00588	.590	.00667	.00490
13	.740	.00836	.00614	.622	.00703	.00516
14	.768	.00868	.00637	.651	.00736	.00540
15	.793	.00896	.00658	.678	.00766	.00563
16	.815	.00921	.00676	.703	.00794	.00583
17	.835	.00944	.00693	.726	.00820	.00603
18	.852	.00963	.00707	.747	.00844	.00620

Fig 8 : REPLICATION DIFFERENCES

(c) BETWEEN-REPLICATION VARIANCE COMPONENT



from these graphs is that genetic drift is a very minor cause of the variability between replications. The major part of this variability must therefore result from the differential action of artificial selection on genetic differences between replications, both those differences which were present after sampling the base population and those which were added later by genetic drift. Scale effects are of course not catered for in these graphs; in the Heavy population they would be operating to increase the discrepancy between expected and observed variability but in the Light population they would reduce it.

1.4 DISCUSSION:

The analysis of this present experiment has contained many problems and this is perhaps ironically its most useful result. When the data from all replications was pooled the patterns of response to upward and downward selection were strikingly smooth (Fig. 1). They accorded well with those expected from additive genetic theory apart from one or two explained fluctuations. But the analyses done here surely show that such superficial conclusions are most unwise. When the Controls were taken into account the results no longer followed additive theory very closely, and when the replications were treated separately they showed considerable fluctuation and heterogeneity in mean weights. The size of this fluctuation and the large variability between replications suggests that the first advice to someone starting a similar experiment would be to carry as many replications as possible, weigh as many pupae in each one as possible, and select more parents in each one. Increasing the population size in this way would decrease genetic drift, sampling and weighing errors, and any errors which arise from assuming that a small population is normally distributed etc. As Kojima (1961) points out, even the presence of gene dominance in a small population could give a big bias to the response.

The second conclusion which impressed the author is that analysis of a selection response is by no means a cut-and-dried affair. Techniques such as fitting a regression line to the response curve should be tested before being blindly used; if they are not so tested a reader should at least be made aware of the statistical assumptions involved and an author must acknowledge that these assumptions could cause considerable bias. Also the results from different replications must not be blindly pooled; in the present experiment some analyses could not be completed to the desired stage because there was too much heterogeneity between the six replications. These two examples should illustrate that analysis of a selection response is very much a developmental process,

whose later stages depend on the results of earlier ones.

The third conclusion which the author regards as important is that neither statistical theory nor genetical theory appear adequate to fully interpret the results of this experiment, though of course this is equally a criticism of the experimental construction itself. The author was unable to completely separate out and measure the many forces affecting his selection response and it was small consolation to know that these forces apparently usually affected the Heavy and Light populations in similar fashion. This is not a wholesale criticism of statistical and genetical theoreticians; it is instead a warning that an experimenter must be prepared to carry large populations and many replications if he expects to assume that this theory will apply to his results. But it does also imply that such theory is poorly adapted to the commercial breeding project, where replication and population size are often both low.

Depressing though these conclusions may sound, there is still considerable value in the techniques of analysis available. The present experiment benefited greatly from the use made of statistical probability theory, and when one test was shown to be invalid or inconclusive there was occasionally an alternative way of looking at the data in an unbiased manner. This approach of analysing selection data by as many methods as possible, inconclusive though each may be on its own, is perhaps the most promising one to take in future. It would at least be an improvement in the sense that (Falconer, 1961; page 208) "..... the drawing of conclusions from the results of experiments in the field of quantitative genetics is to some extent a matter of personal judgment. The further this personal judgment can be pushed back, even to be replaced by something analogous to Decision Function theory, the more reliable will become these analyses of selection response.

PART TWO: PREDICTING SELECTION RESPONSE

In the experiment just described there was very little selection response in either population during Generations 13 to 18, and it would not have been worth the effort as a commercial project. But could this result have been predicted?

Successful prediction of a selection response is perhaps the most-desired ability of plant and animal breeders. Such breeders continually have to decide between various populations, mating systems etc for the project which will bring the biggest reward, and the criteria they have for this decision are "know-how" and selection theory. The author has already discussed in the earlier essay the nature of selection theory that is available. In this next section of the thesis he studies its ability, and validity, to predict a selection response. How successfully it has done so in the past is discussed in the literature review that follows.

2.1 LITERATURE REVIEW:

The literature contains many experiments which attempt to predict the response of a population to selection, then carry out the actual selection programme and compare the realised and predicted responses. The success of these predictions ranges from good to bad but the measure of success is usually a personal judgment of its author. Admittedly he is the person most qualified to make such a judgment and in most cases the present author would not presume to disagree with him. However, sometimes prediction theory may be considered to be inaccurate when in fact the reason for non-agreement with observed response is due to a poorly replicated and badly conducted experiment. It was the present author's intent to retest this prediction theory in an experiment which did not contain these weaknesses.

Notwithstanding that the cause of a poor prediction may sometimes lie with the experiment itself, the present author occasionally finds fault with the judgments of earlier authors. For example Sheldon (1963) says that the results in his Tables 3 and 7 certainly support the prediction that "additive genetic variance would decline and the non-additive genetic variance would increase as selection proceeded" (page 507). But a study of these tables shows no such apparent decline in h^2sire nor h^2dam and no such apparent increase in δ^2_I / δ^2_P, even disregarding that the standard errors are fairly high anyway and also disregarding that δ^2_I / δ^2_P may not be a very useful indicator of the non-additive genetic variance. Similarly the data of Frey and Horner (1955) are interpretable in a manner opposite to that given by the authors. Sometimes the prediction estimates may have standard errors so high as to make them almost meaningless (Falconer, 1953; Kyle and Chapman, 1953). Even recently Kojima and Kelleher (1963) make little importance of this point. Reporting on a selection experiment designed to increase egg production of Drosophila pseudoobscura they calculated the predicted and observed responses for every generation. Then to each of these two responses separately they fitted a regression line over all generations. The two regression lines proved to be nearly parallel. This is indeed good prediction of rate of response (though not necessarily of its absolute amount) but the errors of estimate are not given and the authors admit that they are large. Clayton et al's (1957) prediction studies are also less meaningful than they might appear. Clayton used as the value of "observed" response the mean response of five generations in each of five replications. Although this average value agreed well with the response predicted from the base population the replications themselves varied greatly in observed response, and so prediction for an individual replication was very poor. In the same way Marien (1958) predicted the response of five generations of selection for faster development in Drosophila; his

success was fair but again spoiled by the high variability among his ten replications. The other way in which an apparently successful prediction can be misleading is when there is asymmetry of response present (Falconer, 1953, 1955; Martin and Bell, 1960). In such cases the prediction often agrees with the observed response averaged over both directions but does not agree with either direction separately.

Occasionally some of these prediction tests appear to be invalid, for example in Kojima and Kelleher's paper of 1963. These authors feel entitled to predict negative responses. Kojima explains (pers. comm.) that these predictions arise from estimates of a negative value of additive variance. He argues that although such estimates are undoubtedly due to sampling error they are still valid estimates and so must be accepted; otherwise the average value of all the estimates will be given a positive bias. This argument does not hide the objection that Kojima is surely mis-using genetic theory by accepting negative variances for single-generation predictions. He seemingly should have either re-estimated the variances or else have averaged all of his estimates, those negative and those positive, before predicting from them. The attitude to take to these and other unreasonable estimates has been studied by Harris (1964). In a Monte Carlo simulation programme Harris compared the predicted response to index selection when the negative estimates of additive variance were accepted at face value, with the predicted response when such estimates were converted to zero. He concluded that the difference between them was not important for his selection index model, though curiously the latter system gave a lower and also more accurate prediction.

Lastly, some of the prediction equations themselves are too simplified to be realistic. Consider Griffing's (1960) equation for the theoretical response to truncation selection

$$\mu_n = (i/\sigma^2)\left[n\sigma^2_A + \sum_{r=1}^{n}(1-y)^{r-1}.\tfrac{1}{2}\sigma^2_{AA}\right]$$

where i = selection differential

σ^2 = phenotypic variance

σ^2_A = additive genetic variance

σ^2_{AA} = additive x additive epistatic variance

n = number of generations of selection

y = recombination index.

This equation deals with only two loci but it caters for arbitrary dominance, epistasis, linkage, and number of alleles. It predicts an asymptotic response, because the coefficient to σ^2_{AA} decreases in successive generations. Griffing makes the usual assumptions of small effects by each gene etc and he disregards natural selection, but his equation still greatly over-estimates the response. Gill (1965b) points out the reasons: firstly there is the drift effect which occurs in all finite populations, and secondly there is the fact that σ^2_A would itself be lowered in successive generations. In an experimental test of the equation Gill shows that over-estimation occurs whether prediction is made from base population parameters, from first-generation parameters, or from sixth-generation parameters; and the discrepancy increases as the prediction becomes longer-range. This popular habit of extrapolating a past response to predict a long-term future response is rarely successful. Long (1961) reports on its use during one of the classic plant breeding experiments on long-term selection. The experiment involved 61 generations of selection for oil and protein content in maize and during its progress different authors attempted to predict further progress. They either used a prediction formula or they simply fitted a regression line to the past response and then extrapolated it into the future. Long reports that nine out of the twelve extrapolations were failures, though admittedly nineteen out of the twenty four predictions based on formulae were

also seriously wrong.

This literature review shows how difficult it is to get a perspective of prediction theory. It was therefore decided to seek such a perspective by corresponding with some well-known population geneticists (Appendix 1). Each of these persons was asked, among other things, how well he expected prediction theory to apply to actual populations. This consensus of opinion did not completely clarify the subject but it did make two main points.

(1) Prediction theory mostly has been derived only for "ideal" populations that are in linkage equilibrium, unselected, and random mating; that this theory is valid only for one generation of selection; that even then the limited gene model can give only an approximate prediction.

(2) In most populations such conditions of idealism can never be fulfilled. Therefore prediction theory is strictly not applicable.

Perspective to these points was summed up by Kimura (pers. comm.): "As long as the genotypic variance is largely additive and selection is carried out under random (or nearly random) mating, I should expect that the gain is reasonably well predicted by the current theory, at least for the first few generations, irrespective of the genetic structure of base populations".

The experiment aims to assess this claim and if it fails, to find out why it fails. But even if it is successful, prediction of response from an analysis of the base population means that effort and facilities are being used on strictly non-productive work and that the more productive results of actual selection are being delayed for one generation. If a prediction system could be found which was reliable and yet did not hold up or add cost to the selection process this system would be more useful than others. The author probes this approach by analysing whether the response from the first generation of selection shows any significant correlation with the subsequent mean response over a period, and

indeed any correlation with the standard deviation or regression coefficient of response over the same period. In the event that first-generation response is a poor predictor of subsequent response the author sets up a limited analysis to find out how it may be improved.

2.2 METHODS:

The problems and various methods of predicting a selection response have already been discussed earlier in this thesis. Some of this information will be briefly reviewed before the practical work is described.

Prediction from first-generation response: The simplest method of predicting a selection response is obviously to extrapolate from the response already obtained in past generations of selection. This method will not provide a clear understanding of the type of gene action that causes the response but its value lies in catering automatically for forces such as natural selection that might otherwise ruin the prediction. This method of extrapolation clearly can be applied at any stage of a selection programme and the prediction may be made as long-range as desired (though recognizing the increasing risk of error as the range becomes longer). In the present experiment the prediction is made from the response realised at the first generation of each period, and tested for its closeness to the mean response over the whole six generations of that period. This test involves the use of data that has already been collected for Part One of the thesis, for example the selection responses at Generations 1, 7 and 13 are used to compare with the mean responses respectively of Generations 1 to 6, 7 to 12 and 13 to 18. Once again the analysis is applied to each replication separately but the overall agreement within any one period between mean response and first-generation response is tested by correlating the former with the latter, one pair of items coming from each replication. If the correlation coefficient has a value close to unity, and the average value of first-generation responses is similar to the average value of mean responses, then the first-generation response is considered to be a good predictor of the mean response for any six-generation period. But if the average values of first-generation response and mean response differ

greatly then one is overestimating the other regardless of whether the coefficient is near unity or not. Conversely, if the value of correlation coefficient is not near unity then the first-generation response is obviously a poor predictor of the subsequent mean response, regardless of how similar are their average values. Thus in this way much information can be obtained on the merit of extrapolating from a given selection response in order to predict the further response. If the correlation between first-generation response and mean response is observed to be poor then it would be useful to analyse why this should be so. A fairly crude type of analysis is illustrated, in which the effects of Control and of replication are separated out in turn.

Finally the opportunity is taken to test first-generation response as a predictor of the size of fluctuations between generations and trends between generations. Again this predictability is first explored by the use of correlation coefficients, this time between first-generation response and (separately) standard deviation and regression coefficient, for each period in turn. If any of these coefficients proved significant then the correlation would be studied in more detail, but if these coefficients were not significant there would be no point in taking the analyses any further.

Prediction from base populations: The method described above of predicting future response from past response would appear to be inherently superior to any other method. It automatically caters for effects of scale and natural selection, and it can be carried out under the exact conditions of mating system, selection intensity etc that are to be applied in the succeeding generations. Because of this also, no generations of selection are lost or delayed while the prediction analysis is taking place. The main fault of this method is that it does not provide an understanding of the gene action in the population and without this understanding it is difficult to suggest alternative selection pro-

cedures if no response is being obtained. For example consider a population that contains a high amount of dominance variance but no additive variance. In spite of high genetic variability it will give very little response to mass selection; but it will probably give considerable response to a hybrid mating scheme such as reciprocal recurrent selection, which can make use of non-additive genetic variance. The method given above of predicting from a first-generation response could never provide this type of guidance. This suggests the need for a prediction system which is based on analysing the relative amounts of additive, dominance, epistatic and non-genetic variances in a population. One of the most accurate ways of doing this analysis is by a "diallel cross".

A diallel cross is basically a type of mating experiment in which a random sample of individuals is drawn from a population and then every male in the sample is mated to every female in turn. By measuring and comparing the progeny of all these matings it is possible to estimate the average breeding value (or General Combining Ability) of each male taken over all females, and any uncommonly good nicking effects (Specific Combining Ability) that each male may have with particular females. General and Specific Combining Abilities lead directly to an estimate of the additive and dominance variances respectively in the population and these in turn lead to a prediction of the heritability that should result from selection in the population. The actual response which may be expected from selection is then obtained merely from multiplying this estimate of heritability by the selection differential.

The earlier essay on selection theory explained how the genetic analysis of a population, as for example by a diallel cross, depended on partitioning the phenotypic variance into between-family and within-family components. These observational components are themselves compounded from causal genetic components, that is from known fractions of the additive, dominance and epistatic variances.

Separation of these components depends upon being able to observe the variances of several types of families, such as halfsib families, fullsib families, cousins etc. But if for some reason the genetic analysis must be restricted in scope it is often not possible to separate out all these components of genetic variance, and so an estimate of any one component will probably be biased upwards by small fractions of the other components. In the present prediction experiment such restriction on the scope of genetic analysis was considered to be wise, though it meant that most of the estimates of heritability would be biased upwards, especially those from fullsib covariance analysis. The reason that such restriction was made on the scope of analysis was that the author considered that the Tribolium stocks he was using would be unable to complete the programme of a full diallel cross. This diallel cross was to be carried out both on the base population (Generation 0) and on the Generation 18 Heavy and Light populations as this would indicate any changes in the variance components that were caused by selection. However, it soon became obvious that the Light replications would be very small and weak by Generation 18 and consequently there would be too much risk of failure if they were committed to a long and complicated diallel mating programme. Accordingly the author adopted a restricted form of diallel analysis which uses a 2 sire - 2 dam block mating system, as described in Sheldon (1963).

Within each 2 sire - 2 dam block the parents are mated in pairs, then they are interchanged. This system requires that they be left together long enough in the first arrangement to produce enough progeny for measuring, and that after these parents have been rearranged the second progeny be not sampled until it is sure that the pupae present have indeed been sired by the second male and not by residual sperms of the first one. This data on timing procedure was not available from the literature and so a pilot experiment was set up to provide it. The details of this pilot experiment will not be given but it made the following

points:

(1) Pair-matings required up to a fortnight to produce enough progeny for weighing needs, especially so in the Light population.

(2) In about 30% of the mating blocks one of the four parents died before the second progeny could be sampled. This failure rate was again highest in the Light population.

Based on this pilot information the diallel analysis experiment was set up as follows:

(1) The analysis was carried out on a base population and on the Heavy and Light populations that resulted from selection in the thesis. The base population used for this diallel analysis was a mixture of the Heavy and Light stocks brought from Ruakura, that is those populations from which the replications were sampled at Generation 0 in the thesis. These original two base populations were allowed to mass-mate separately until selection was finished at Generation 18, then equal numbers of mating blocks were set up from each of them and these represent the "base population" in Table 19.

(2) The replications of Generation 18 selection lines were analysed independently of one another but equal numbers of mating blocks were set up from each one in the Light and Heavy populations. Replication 8 of the former was very weak however and so was replaced by Replication 7 before the diallel analysis was started.

(3) The analysis carried as many mating blocks as facilities would allow. More blocks were set up for the Light population than for either the Heavy or base populations, because of its greater rate of failure in the pilot experiment.

(4) The parents for each mating block were chosen randomly from a replication. Perhaps more precise data would have been obtained if they were mated assortatively (by virtue of increasing the midparent variance). However, such estimates are open to bias in small populations if large non-additive effects are present (Reeve, 1961). Moreover, Dawson (1964) showed for Tribolium that because maternal effects were present in his population, assortative mating inflated the regression value of offspring on parent.

(5) Parents were left to mate for 14 days then remated to the other partner for

another 28 days. The second progeny were destroyed until 14 days after the re-mating: this ensured that those pupae eventually weighed had indeed been sired by the second male (Tribolium Information Bulletin, 1960).

(6) Six female and six male pupae were weighed in each progeny.

The diallel cross described above was designed to provide estimates of additive genetic variance, and thence of heritability, in each of the three populations. Within each population the data could be manipulated to provide three or four fairly independent estimates of the heritability, according to the methods listed below.

(a) Halfsib covariance analysis
(b) Fullsib covariance analysis
(c) Parent-offspring regression analysis
(d) Abplanalp's linear heritability analysis.

The relative merits of the first three methods have been discussed earlier in the essay on selection theory and it will suffice to repeat here merely that the amount of bias from non-additive variance is least in regression analysis, intermediate in halfsib analysis, and greatest in fullsib analysis. The fourth system is based on analysis of displacement data rather than on analysis of variance data and it is potentially superior to all other systems in dealing with the actual prospective parents and thereby in catering for asymmetrical responses to selection.

The halfsib analysis involved the data from all 2 sire - 2 dam blocks which completed their mating programme and produced the required number of progeny pupae for weighing. The fullsib analysis involved separate data entirely, taken from those mating blocks which completed the first part of their programme but not the second; also data from the pilot experiment. The parent-offspring regression analysis also used data from only the first matings of each 2 sire - 2 dam block. On this occasion the female and male progeny data were pooled and regressed against dam weight, then against sire weight. Abplanalp's linear

heritability analysis was applied only to the base population and not to the Generation 18 Light and Heavy populations. The reason for this decision was that such a system which caters for uni-directional selection must be safeguarded against response fluctuations caused by a fluctuating environment. The system is therefore better used to predict an average response over a period (whereby such fluctuations should cancel out) rather than over a single generation. Because of this the system was not used on the Generation 18 populations - there would be no succeeding period of generations by which to test its prediction. The way in which this linear heritability analysis was adapted to the diallel data needs explaining. Only half of the data in each 2 sire - 2 dam block were used, namely from those progeny which were sired by one of the two males. The progeny which were sired by the second male are in no way involved in the analysis. In this way the data fit into a hierarchial analysis, effectively each sire being mated to two random females and each female producing six female and six male pupae for weighing purposes. The derivation of this heritability estimate is given in detail by Abplanalp and is simple to follow, but it is too long to warrant repeating in this thesis. The selection intensity used in the analysis corresponded to that used later in the experiment, namely that in this base population the heaviest 4/32 of progeny were treated as prospective parents for upward selection and the lightest 4/32 were treated as prospective parents for downward selection. The heritability calculations then exactly follow the pattern of Abplanalp's model, though he states that it is difficult to assess the error of such a heritability estimate.

Table 19: Estimation of Heritability

Base population

Half-sib analysis

	d.f.	Mean Square	d.f.	Mean Square	
				female progeny	male progeny
Between blocks	L-1	$\sigma^2_W + 6\sigma^2_I + 12\sigma^2_{S(or\,J)} + 24\sigma^2_D$	36	.12028	.10667
Sires within blocks	L	$\sigma^2_W + 6\sigma^2_I + 12\sigma^2_S$	37	.10757	.05486
Dams within blocks	L	$\sigma^2_W + 6\sigma^2_I + 12\sigma^2_D$	37	.06486	.07514
Sire X Dam interaction	L	$\sigma^2_W + 6\sigma^2_I$	37	.05243	.04027
Residual	4(6-1)L	σ^2_W	740	.04927	.02723

$$h^2_{sire} = \frac{4\sigma^2_S}{\sigma^2_W + \sigma^2_I + \sigma^2_J + \sigma^2_S + \sigma^2_B}$$

h^2_{sire} =	.32 ± .10	.14 ± .08
h^2_{dam} =	.07 ± .07	.33 ± .10

Full-sib analysis

Between progenies	M-1	$\sigma^2_W + 6\sigma^2_P$	41	.22220	.15390
Within progenies	5M	σ^2_W	210	.05610	.05010

$$h^2 = \frac{2\sigma^2_P}{\sigma^2_W + \sigma^2_P}$$

h^2 =	.66 ± .14	.51 ± .14

Regression analysis (74 progenies, each of 6 females and 6 males)

$h^2 = 2b_{P\bar{O}}$

regression on dam	h^2 =	.26 ± .05	
regression on sire	h^2 =	.04 ± .04	

Linear heritability analysis (55 pupae chosen as prospective parents from the 444 weighed)

$$h^2 = \frac{2(\Delta S + \Delta D)}{\Delta S + \Delta D + \Delta F}$$

		female	male
h^2_{upward}	=	.18	.41
$h^2_{downward}$	=	.10	.29

Realised heritability : **Generation 1** (average of 12 reps, each of 32 pupae)

$$h^2 = \frac{\text{selection response}}{\text{selection differential}}$$

h^2_{upward}	=	.40	.45
$h^2_{downward}$	=	.02	-.04
($h^2_{average}$	=	.21	.20)

: **Generations 1 to 6**

h^2_{upward}	=	.28	.27
$h^2_{downward}$	=	.04	.04
($h^2_{average}$	=	.16	.15)

Table 19: (Continued)

		Heavy population (Gen 18)			Light population (Gen 18)	
Half-sib analysis		female progeny	male progeny		female progeny	male progeny
Between blocks	30	1.54733	1.10167	22	.33864	.28091
Sires within blocks	31	.26000	.28645	23	.05261	.04478
Dams within blocks	31	.27323	.29032	23	.07435	.06087
Sire X Dam interaction	31	.10839	.16323	23	.10565	.08783
Residual	620	.07269	.07469	460	.04274	.03572
h^2_{sire} =		.32 ± .10	.28 ± .10		-.31 ± .01	-.30 ± .14
h^2_{dam} =		.35 ± .11	.29 ± .11		-.18 ± .04	-.19 ± .04
Full-sib analysis						
Between progenies	92	.50435	.45196	58	.27810	.25155
Within progenies	465	.05882	.05465	295	.05020	.04481
h^2 =		1.12 ± .09	1.10 ± .09		.86 ± .12	.87 ± .08
Regression analysis		(62 progenies)			(46 progenies)	
regr. on dam h^2 =		.78 ± .07			.28 ± .08	
regr. on sire h^2 =		.47 ± .07			.27 ± .08	
Realised heritability		(6 reps)			(6 reps)	
: Generation 18		female	male		female	male
h^2_{upward} =		-.91	-.81		-.32	-.08
$h^2_{downward}$ =		.64	.68		.27	.10
($h^2_{average}$ =		-.13	-.06)		(-.02	.01)
: Generations 13 to 18						
h^2_{upward} =		-.14	-.13		.02	.07
$h^2_{downward}$ =		.24	.15		.06	.03
($h^2_{average}$ =		.05	.01)		(.04	.05)

2.3 RESULTS:

Prediction from base population: The various estimates of heritability are given in Table 19. Also included in the table are the formulae used in the estimation, and the values of mean square and degrees of freedom used in the formulae. This extra information is included on the advice of Hanson (1963): he considers that raw heritability estimates should have little meaning to other readers if the supporting data is not given. This situation arises because the term "heritability" has been used in many ways and not all of them are equivalent; therefore it is unwise to compare or extrapolate heritability estimates without knowing how they were derived. Table 19 also contains the realised heritability results for Generations 1 and 18 and for the means of Generations 1 to 6 and 13 to 18, these results acting as the "observed" response with which to compare the above estimates.

The only other comments to make before assessing the results are firstly that in the halfsib analysis the estimate of σ^2_B (between-blocks component of variance) can be obtained from subtraction of either the sires-within-blocks mean square or the dams-within-blocks mean square. For the purpose of this experiment σ^2_B was estimated both ways and then averaged for use in the heritability formula. Secondly the values given for realised heritability at Generation 1 and Generations 1 to 6 are obtained by averaging the results from both Heavy and Light populations. For example, h^2upward in the table is the mean value for h^2 in the main line of the Heavy population and the Control line of the Light population; and h^2downward is the mean value for h^2 in the main line of the Light population and the Control line of the Heavy population.

The following comments can be made about Table 19.

(1) Realised heritability: There is a fall of value between Generation 1

in the base population and Generation 18 in the Light and Heavy populations, from an average of about $h^2 = .2$ to $h^2 = -.05$. A similar fall (.16 to .04) is present between Generations 1 to 6 in the base population and Generations 13 to 18 in the Light and Heavy populations. This combined evidence is sufficient to accept that selection has caused the heritability to fall appreciably during the 18 generations. This pattern of change now is used as the standard with which to compare the estimates of heritability predicted by the various methods in the table.

(2) Halfsib analysis: Taking into account their standard errors, the four halfsib estimates within each population are fairly consistent with one another. In the base population the mean of these four values is very similar to that of the realised heritability data and so it may be accepted (tentatively because of the size of standard error) that halfsib analysis is a satisfactory method of predicting heritability response from such a base population. However, the similarity between halfsib estimates and realised heritability is much less at the end of selection. Whereas the latter estimate is close to zero in both Heavy and Light populations the former predicts heritability values of about + .30 and -.25 respectively. By taking the same attitude to the negative value as he took to Kojima and Kelleher's (1963) negative estimates the author would substitute a predicted value of $h^2 = 0$ in the latter case, and this now accords with the realised value. But in the Heavy population the closeness of all four estimates to the value $h^2 = .30$ must be considered to overrule the excuse that standard errors are quite large, and it must be accepted that halfsib analysis has not successfully predicted the realised heritability in this population. The possible reasons for this adjudged failure will be considered shortly in relation to the other systems of prediction that follow.

(3) Fullsib analysis: Fullsib estimates of predicted heritability are much too high in the base population and they actually increase from this to the

Light and Heavy populations. Neither in absolute value for any one population nor in pattern of change from Generation 1 to Generation 18 do they resemble the realised heritability data. On this evidence it must be accepted that full-sib analysis overestimates the true heritability by such a large amount as to make it virtually worthless for predicting the quantitative response.

(4) Parent-offspring regression analysis: There is reasonable agreement in the base population between the heritability values predicted by this method and the values realised. But once again the values predicted from the analyses at Generation 18 are too high, and by an amount that makes them unreliable for prediction purposes.

(5) In the Drosophila experiment of Clayton et al., (1957) the base population estimates of predicted heritability were obtained by halfsib analysis, fullsib analysis, and parent-offspring regression analysis; and they gave very similar estimates to one another. In the present experiment the estimates from fullsib analysis are consistently higher than the others. This upward bias could be due to any one of a number of factors, such as dominance and epistatic variance, and variance due to environmental and maternal effects common to full-sib litters etc. Particularly is there evidence for maternal effects. It might perhaps be considered that the egg-laying habit of Tribolium would minimise the amount of non-genetic influence that a dam could have upon her progeny, but in fact considerable maternal effects are recorded by Blair (1962) for the character of larval weight in Tribolium and by Dawson (1965) for the character of developmental rate. There is evidence of a similar maternal effect for the character of pupal weight in the present experiment. This evidence is the observation that in all three populations analysed, the coefficient of regression of progeny weight on dam weight was greater in value than that of progeny weight regressed on sire weight.

Sheldon (1963) would probably consider of importance the fact that sire x

dam interaction variance increases greatly relative to the residual variance. He regards this ratio of variances as a good indicator of the amount of non-additive genetic variance present in the population and its rise in value from Generation 0 in the base population to Generation 18 in the Light and Heavy populations would explain partly the parallel rise in predicted estimate of heritability, since all three methods of prediction contain non-additive genetic variance as a source of upward bias. However, it should be remembered that "non-additive" includes many factors and interactions besides dominance and epistasis. Miller et al., (1963) claim that the ratio of interaction variance to phenotypic variance is not after all a very useful indicator of the non-additive genetic variance. For example it includes even the interaction between maternal environment of the parent and genotype of the progeny.

Scale effects are not expected to be a major cause of any discrepancy between predicted and realised heritabilities. Falconer (1961) explains that scale has very little influence on realised heritability because it affects the numerator (selection response) and denominator (selection differential) by almost equal amounts. However, a simple mathematical argument shows that scale effects will bias upwards the estimate of heritability from parent-offspring regression if the mean weight of offspring exceeds the mean weight of parents.

This argument proceeds as follows (supplied by Professor A.L. Rae):

Assume X_i = observation on i th parent
Y_i = observation on offspring of i th parent
Let μ_p be mean value of parents
g_{pi} be deviation from mean due to additive gene action
ε_{pi} be random deviation
and let μ_o, g_{oi}, ε_{oi} have similar meaning for offspring.

Then if effects both genetic and environmental are multiplicative it can be written that $X_i = e^{(\mu_p + g_{pi} + \varepsilon_{pi})}$

$= e^{\mu_p} \cdot e^{(g_{pi} + \varepsilon_{pi})}$

Expand in a Maclaurin series

$$X_i = e^{\mu_p}\left[1 + (g_{pi} + \varepsilon_{pi}) + (g_{pi} + \varepsilon_{pi})^2 \Big/ 2! + \ldots\right]$$

Neglecting quadratic and higher terms, and writing

e^{μ_p} as M_p gives

$$X_i = M_p + M_p g_{pi} + M_p \varepsilon_{pi}$$

Similarly $Y_i = M_o + M_o g_{oi} + M_o \varepsilon_{oi}$

Then E (regression of offspring on parent)

$$= \frac{\frac{1}{2} M_p M_o Z_g^2}{M_p^2 (Z_g^2 + Z_\varepsilon^2)}$$

$$= \frac{\frac{1}{2} M_o (Z_g^2}{M_p (Z_g^2 + Z_\varepsilon^2)}$$

Thus a bias is introduced if the mean weights of parent and offspring differ. These mean weights for the three populations under analysis are given below.

	Base pop. (Gen 0)	Heavy pop. (Gen 18)	Light pop. (Gen 18)
Parent mean weight (mgms)			
Female	2.16	2.74	1.44
Male	2.07	2.57	1.36
Progeny mean weight	2.09	3.20	1.66

In the base population the progeny mean does not differ greatly from either of the parent means and so no bias from scale effects would be expected. But in the Heavy and Light populations the progeny mean is consistently and markedly greater than the parent means, and as shown above this would tend to considerably inflate the predicted heritability. This factor therefore is tentatively accepted as being partly responsible for the over-high heritabilities

predicted from regression analysis at Generation 18.

Non-additive genetic variance, maternal effects and scale effects have been shown as likely causes of bias in predicting the heritability at Generation 18 and it is likely that there are other causes as well. This total evidence suggests that it is very risky to undertake such prediction by variance analysis in any population that has undergone selection for a long time. Too many new pressures and effects build up which cannot be catered for in such an analysis.

(6) Linear heritability analysis: This method has successfully predicted that there will be an asymmetrical response from the base population. The absolute values of this predicted response are in the region of those realised without being exactly similar to them. As explained earlier the author deemed that such analysis should not be repeated in the Light and Heavy populations but it is obvious that the extremely significant between-blocks mean squares at Generation 18 would bias upwards the linear heritability estimates by very large amounts. As it operates in the base population however, the linear heritability analysis is a better method of predicting uni-directional response to selection than is the method of variance analysis. This is contrary to the conclusion reached by Bhat (1965), in whose experiment the linear heritability estimate exceeded both the variance analysis estimate and the realised heritability by a considerable amount.

The success achieved at predicting heritability in this experiment varied considerably with the method used and generation tested, but on the whole it was probably not quite as successful as the predictions reported by Dawson (1965). Working with a similar selection character, developmental rate in Tribolium, Dawson applied some of the methods used here and some not used here, the latter usually involving inbred lines. He found that predictions were often accurate over short-term periods (7 generations) but mostly inaccurate over longer

periods (13 generations). A main cause of this inaccuracy was considered to be large amounts of dominance variance in the population. The presence of non-additive variance was also stated by Sheldon (1963) to be a primary reason for the predictions in his own work with Drosophila to be unreliable. Sheldon suggested in his final discussion that for a one-way selection experiment, a generation of actual selection may itself be the best predictor of future response. The present author also takes this attitude, and the following pages deal with the ability of a first-generation response to predict subsequent responses. This approach is particularly worthwhile in the present case when it is remembered that in the section above, the predicted heritability was tested against realised heritability averaged over all six replications, and that this predictability would be even poorer than it actually was if the replications had been considered individually. In the section below they are in effect considered individually and so this is a much more meaningful test of predictability.

Selection differential: To be able to predict the heritability of a selection response is of course only halfway toward predicting the response itself. It is necessary also to either predict or measure the selection differential. Fortunately a breeder is usually able to actually measure the selection differential rather than predict it, because in the process of measuring the whole population in order to choose parents he automatically collects the data required to measure the differential. There are a few occasions when the differential will have to be predicted; for example when a breeder does not have time to measure the whole population before deciding if he will get a worthwhile absolute response from selection. In this case he would have to measure a small sample of the population, estimate the standard deviation of this sample (and thereby of the whole population), and then predict the value of selection differential from the formula

$$s = \bar{\iota}\sigma_p$$

where $\bar{\iota}$ = selection intensity in standard units

σ_p = phenotypic standard deviation of sample

The sources of error in this prediction formula are twofold; firstly in estimating $\bar{\iota}$, secondly in estimating σ_p and assuming it relates to a normal distribution. Both of these errors are reduced as the size of sample and number of parents selected are increased. However, the present experiment cannot contribute to measuring these errors as sources of poor predictability for the reason that the sample itself has been used both to represent the population in mean weight and to provide the future parents. Hence no further study of the selection differential has been made in this thesis - on all occasions it could be directly measured rather than predicted.

Prediction from first-generation response: The accuracy of first-generation selection response in predicting the response of other population parameters will be measured in terms of the correlation coefficient between the first-generation response and these other parameters (separately). It has already been explained how this method is a satisfactory first step only, and how the analysis would have to be continued in a more refined way if these correlation coefficients proved to be high in value. The parameters which were compared with first-generation response in this way were response mean, standard deviation, and regression coefficient, for each 6-generation period in turn. Each correlation coefficient is calculated from six sets of paired values (one per replication); for example the correlation with mean response of a given period involved pairing together the first-generation response (x) of that period with its mean response (y), for each replication in turn. Then the correlation coefficient (r) is calculated as

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{(x-\bar{x})^2(y-\bar{y})^2}}$$

The significance of this value of coefficient was determined from Appendix 4 of Bailey (1959) for (n-2) = 4 degrees of freedom. The .05 (.01) levels of significance are r = .81 (.92) and these values may be compared with the actual values of r in Table 20. Also contained in the table are the average responses of first-generation and period mean. For all data in the table, Control values have been averaged into the first-generation response (as well as into mean response) otherwise the effect of a fluctuation in the environment could influence the results too much over a single generation of response.

Once again the correlation between first-generation and mean response has been estimated for both ways of calculating this mean response, namely from linear regression and from the total response divided by six.

Table 20: Correlation between first-generation response and other parameters

Generations	1 to 6			7 to 12			13 to 18		
	1st gen.	mean	r	1st gen.	mean	r	1st gen	mean	r
				Heavy population					
r (first : mean - by regression)	.0729	.0576	-.04	.0687	.0629	.78	.0318	.0164	-.35
r (first : mean - by total/6)	.0729	.0601	-.22	.0687	.0620	.70	.0318	.0172	-.17
r (first : standard deviation)			-.08			-.06			-.50
r (first : regression coefficient)			-.51			-.28			-.27
				Light population					
r (first : mean - by regression)	.0514	.0447	.82	.0446	.0474	.40	.0321	.0135	.57
r (first : mean - by total/6)	.0514	.0433	.94	.0446	.0494	.30	.0321	.0135	.34
r (first : standard deviation)			.03			-.03			.46
r (first : regression coefficient)			-.65			-.54			-.61

On the whole there is very low correlation between the first-generation and mean responses of any period regardless of which method is used. Only once is the value of coefficient significant and twice it is even negative. This evidence is not to say that no such correlation exists but it does suggest that many more paired values than six would be needed to prove such a correlation. The second point to observe from the table is that on all but one occasion the first-generation response exceeds that of the mean. This consistent bias is further evidence that within each period the rate of response gradually falls from first generation to sixth, as was shown earlier by the negative signs of regression coefficients in Table 14.

The correlation between first-generation response and standard deviation is neither large nor consistent. It is not worthwhile analysing the relationship any further than this.

The correlation between first-generation response and regression coefficient consistently has a negative sign. This means that the higher is the response at the start of a period the faster it falls off during the period. This conclusion accords with genetic theory which states that the amount of genetic variability in a population affects in the same way both the amount of response obtained from selection and the rate at which this response falls off. However, the absolute values of these negative coefficients are too varied and too far below the significance levels of .81 (.92) to have any predictive value.

The total impression of results in Table 20 is depressing in some ways, particularly in its evidence that even when the Control variation is taken into account the selection response at any one generation is a poor predictor of future response. There could be many reasons for these results but such reasons fit into two main camps - either there <u>is</u> a good correlation between

first-generation response and later response but the present experiment could not detect it, or alternatively no such correlation exists. Additive genetic theory confirms that later response will be less than first-generation response, as it was in Table 20, but this bias should not affect the correlation coefficient. Since genetic trends therefore are apparently not to blame for the poor correlation (and are very small when compared with the fluctuations anyway) it appears likely that fluctuations of response may be responsible. Furthermore, since these fluctuations are themselves caused by environment, genetic drift and sampling error etc, which are all factors able to be controlled by the experimenter, it is likely that the experiment itself was the cause of the low correlations obtained. Perhaps there were too few replications, perhaps each replication had too few animals. This hypothesis is investigated in the analysis that follows, by separating out some causes of fluctuation and measuring their effects on the results. It is admitted that the analysis is very limited but it does give the author some understanding of the poor correlations.

The method of analysis is illustrated in detail for the response period of Generations 1 to 6 in the Heavy population. In this period there are the usual six replications and each contains an upward selection line and a downward Control line, that is twelve lines altogether. For each of these twelve lines the author calculated the first-generation response and the mean response over all six generations; then calculated the discrepancy between the two. The values of these discrepancies are given below (in mgm per pupae).

	Rep 1	2	3	4	5	6
selection upwards	-.14	+.02	+.30	-.04	+.04	+.17
downward (Control)	+.12	-.07	-.12	+.04	-.08	-.07

The average size of these deviations (disregarding the sign) is .102 and their standard deviation is .130 (which is a measure of their range). Since the

mean response during this same period was only .0601, these measures of deviation are both very large indeed. In this sense the first-generation response is a very poor predictor of the mean. But suppose now that the responses of an upward line and its Control are averaged before the deviations are calculated.

The new deviations are

Rep 1	2	3	4	5	6
-.01	-.02	+.09	.00	-.02	+.05

and the average size of them is .031 (their standard deviation equals .044). This is now a much smaller percentage of the mean and the author contends that this result can be interpreted thus: there is a very considerable improvement in predictability if environmental sources of error are catered for by an adequate Control.

Suppose now instead that the upward and downward (Control) lines are left separate but that within each one the responses of all six replications are averaged, before the deviations are calculated. These deviations would be

upward	+.06
downward	-.03

and the average deviation is .045 (standard deviation equals .062). This is again a great reduction on the original discrepancy above, though slightly less improvement than by the correction for Control. This result suggests to the author that correlation between first-generation and mean responses would probably have been higher if larger populations had been used, though it must be conceded that the pooling of six replications for this analysis above is not necessarily equivalent to a single population six times the size of each one.

Finally, suppose that the responses of all twelve lines are averaged before the deviation is calculated between first-generation and mean. This deviation

measures .013, that is less than 25% of the mean.

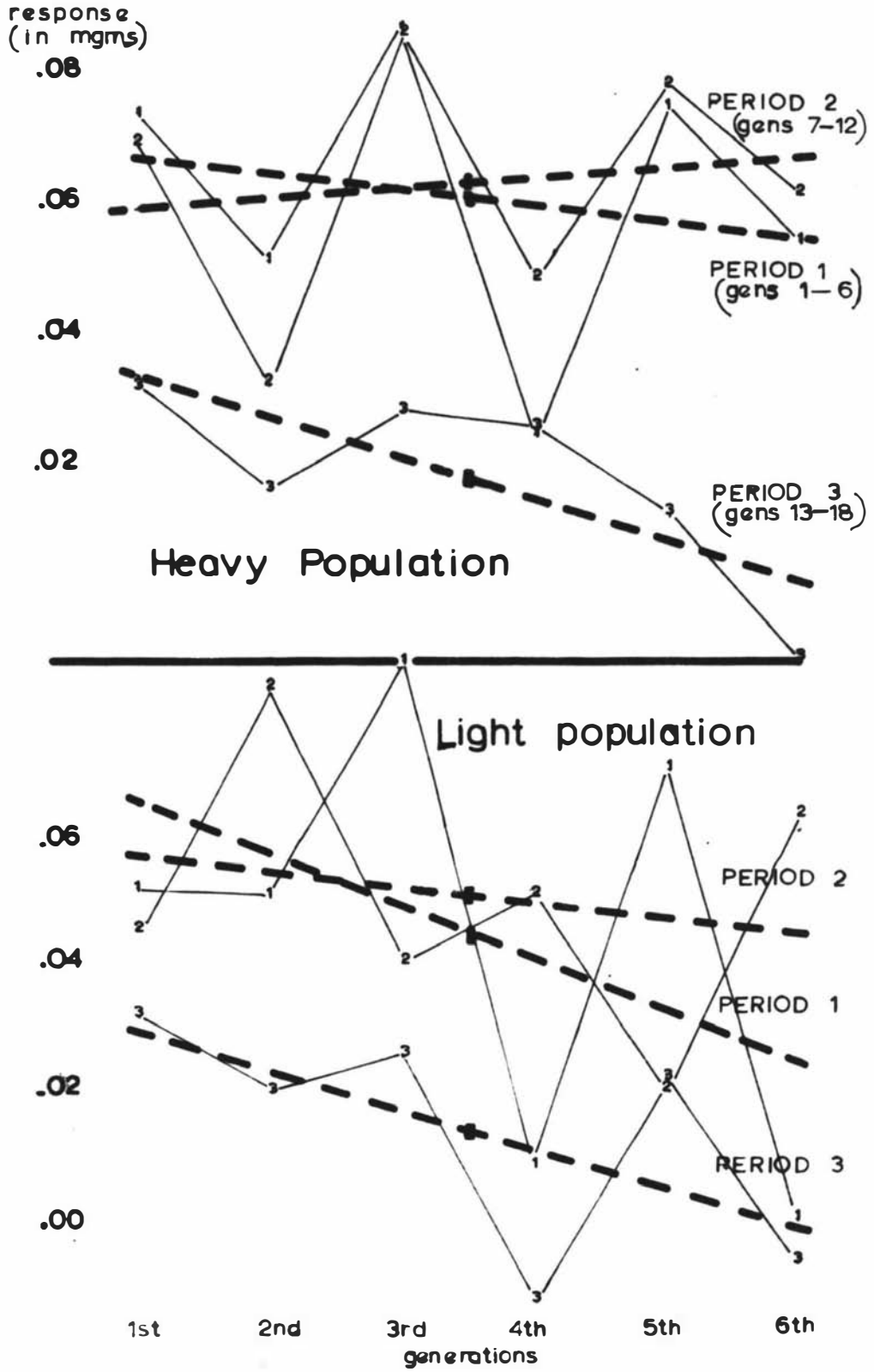
The author has given his interpretation of the analysis above but before any general conclusions were made the same analysis was applied to each of the remaining periods of selection. The results are contained in Table 21, and include the value of period mean response, average discrepancy of this mean from period first-generation response, and the standard deviation of these discrepancies.

Table 21: Discrepancy between first-generation response and mean response

Generations	1 to 6		7 to 12		13 to 18	
			Heavy population			
mean response	.0601		.0620		.0172	
	average discrep.	S.D.	average discrep.	S.D.	average discrep.	S.D.
Control and reps both unpooled	.102	.130	.103	.127	.075	.098
Control pooled, reps unpooled	.031	.044	.022	.031	.033	.037
Control unpooled, reps pooled	.045	.062	.021	.030	.047	.068
Control and reps both pooled	.013		.007		.015	
			Light population			
mean response	.0438		.0494		.0136	
	average discrep.	S.D.	average discrep.	S.D.	average discrep.	S.D.
Control and reps both unpooled	.055	.070	.096	.133	.042	.050
Control pooled, reps unpooled	.013	.015	.062	.083	.031	.031
Control unpooled, reps pooled	.012	.018	.008	.011	.023	.033
Control and reps both pooled	.009		.002		.019	

In all cases the average discrepancy and its standard deviation are greater than the mean itself if the replications are treated separately (unpooled) and their Controls are also treated separately from the main selection lines. These two parameters reduce considerably when Control data is incorporated (pooled) beforehand, and equally so when the six replications in each direction are pooled beforehand. If both of these tasks are carried out beforehand then there is even further improvement. The average discrepancy which remained at this last stage varied from 4% to 136% of the mean responses and this remainder is undoubtedly caused partly by genetic trends and partly by fluctuations which have not been completely catered for. The relative importance of these two forces, directional and non-directional, is given in Figure 9 in which the regression line is shown fitted to the graph of response values. Sometimes the line slopes from the period mean towards the first-generation value, for example in the period covering Generations 1 to 6 of the Heavy population. In such cases the deviation between first-generation and mean is caused partly by trends of falling response. But at other times the regression line slopes away from the first-generation value (for example in period 2) and in these cases the deviation cannot be explained by such trends. It is obvious that where the slope of regression line does not explain all deviations, then these deviations are apparently caused by fluctuations of response. This means that despite the pooling of responses from all six replications and Controls the population carried was not big enough to make these fluctuations unimportant.

Fig 9: FLUCTUATIONS AND TRENDS OF RESPONSE



Prediction of limit response: The final analysis undertaken in this section of the thesis was not planned for at the start but was made possible by the subsequent response to selection. A study of the response curves in Figure 1 shows that there was very little selection progress made in either population during the final few generations, and so it was tentatively accepted that the limit of selection response had been reached. The opportunity was therefore taken to test a further piece of prediction theory involving the first-generation response. This prediction theory is based on additive gene action and as presented by Robertson (1960) it estimates that the total amount of selection response obtained when the limit is reached measures about $2N_e$ times the amount of first-generation response. The prediction is only approximate and also applies only when the product of N_e (effective population size) and i (selection intensity) is low, but unfortunately Robertson does not specify how low. In the test of this theory carried out below the value of limit response was taken as the highest generation mean reached by each Heavy replication during the 18 generations of selection (the lowest generation mean in the case of Light replications). That this value was a fair one to take was accepted only after a study of Figure 1 Controls showed that none of these generation means was apparently due to a serious fluctuation of environment. The "total response" was then obtained by subtracting from this limit value the replication mean at the beginning of each period, that is the replication means of Generations 0, 6 and 12. In each of the three cases, first-generation response was again measured as the average of a main selection line and its Control, in order to cater for environmental fluctuations. The ratio of total response to first-generation response is given in Table 22 for the 18, 12 and 6 generations of selection involved. According to Robertson's theory this ratio should have a value of $2N_e$:1, and since N_e has already been estimated as 4.7 in the Heavy population and 6.4 in the Light population the predicted ratios are respectively

9.5:1 and 12.8:1.

In Table 22 the values of total response and first-generation response have been rounded off after the ratio was calculated.

Table 22: Ratio of total response to first-generation response

Generations		1 to 18			7 to 18			13 to 18		
		total resp.	1st-gen resp.	ratio	total resp.	1st-gen resp.	ratio	total resp.	1st-gen resp.	ratio
					Heavy population					
Replication	1	1.02	.07	15.1	.38	.06	6.2	.07	.03	2.6
	2	.76	.03	28.1	.40	.04	9.5	.15	.05	3.0
	3	1.33	.12	10.9	.82	.12	6.9	.06	.02	2.6
	4	1.46	.08	17.7	.70	.09	7.4	.29	.01	24.6
	5	1.05	.04	26.2	.41	.07	5.9	-.01	.05	-0.1
	6	1.14	.10	11.1	.47	.02	23.5	.14	.02	6.6
average ratio				18.2			9.9			6.5(7.9)
predicted ratio				9.5			9.5			9.5
					Light population					
Replication	1	.36	.03	13.1	.46	.02	18.6	.19	.05	3.7
	2	.45	.06	7.8	.32	-.01	-21.3	.10	.07	1.6
	3	.13	.05	2.7	.11	.04	2.6	.03	.04	0.9
	4	.37	.08	4.5	.27	.05	5.2	.14	.07	20.7
	5(or 8)	.48	.08	5.6	.29	-.04	-7.8	.07	.05	1.4
	6	.15	.01	15.0	.26	.21	1.3	.14	-.02	-5.6
average ratio				8.1			-0.2(6.9)			3.8(5.7)
predicted ratio				12.8			12.8			12.8

On three occasions the average ratio has been given an alternative value in brackets. This alternative represents the value obtained when all negative ratios are omitted from the average. The author considered that since each first-generation response represents an average value of selection in two directions then a negative value is theoretically impossible and must be due to weighing or sampling error and could therefore be omitted. However, since the remaining ratios might otherwise be claimed to receive a positive bias the average ratio has been calculated by both methods.

It is interesting to note, but difficult to explain why, the average ratio decreases as the selection programme nears its end. However, the result which is considered to be most important is that realised and predicted ratios are of a reasonably similar order throughout the table. A possible reason for the realised ratio to be greater than predicted in the Heavy population and less than predicted in the Light population, lies in the effect of asymmetry on the method of calculating first-generation response. In both populations this asymmetry effect is maintained in the main selection lines but by incorporating the Control it is cancelled out of the first-generation response; therefore it biases the measurement of "total response" but not that of "first-generation response". The effect of this bias is to increase the ratio above predicted value in the Heavy population and decrease it below predicted value in the Light population, exactly as the table shows.

2.4 DISCUSSION:

In the present experiment there was reasonable agreement between the realised heritabilities and those predicted from halfsib analysis or from regression of progeny on sire, but this agreement was seldom precise enough to use with confidence in future experiments, especially when the size of standard error was taken into account. Those predictions which were well astray, namely from fullsib analysis and from regression of offspring on dam, could usually be explained with quite reliable evidence of non-additive genetic action. It could be argued that if a breeder was to know beforehand that such non-additive variance was present in a population, then he could make mental compensation for the upward bias that it causes, and thereby still predict the heritability with reasonable accuracy. But this is a matter to be determined by the gambling instincts of the breeder.

The results of this second section of the thesis differ slightly from those of Sheldon (1963) but the general conclusions would be about similar. Referring to his efforts to predict selection response from a 2 sire - 2 dam block system of diallel analysis Sheldon concludes that - "In general, neither the base population nor later parameters predicted adequately the subsequent selection responses" (page 490).

It seems reasonable to conclude that the standard methods of predicting heritability of response are not precisely accurate, even in a base population, mostly because these methods cannot cater for all of the non-additive genetic properties of a population. For example they cannot cater for asymmetrical responses. This suggests that the prediction systems which would be most likely to succeed would be those ones relating most closely to the actual processes of selection. Abplanalp's system of linear heritability analysis offers considerable promise in theory and achieved considerable success in practice; it

certainly warrants more test by experiment. The other method tried in this thesis, which makes use of the response realised from one generation of selection, also ought to be successful in theory. It is based on the theory that this first-generation response depends on the genetic variance already present and also on various non-genetic forces which might be operating; and that similar genetic variance (though slightly reduced) and non-genetic forces might be expected to continue to be present and operating in future generations. Therefore a correlation might be expected between first-generation response and subsequent mean response. The fact that this correlation proved to be not very significant in the present experiment is more a criticism of the practical aspects than of the theory. Analyses of various results by the author indicated that such correlation would possibly have been much higher if more replications had been carried and more animals weighed in each one. In other words, sampling error in all generations was as much a cause of poor predictability as was the first generation itself. It was the author's intention that if a high correlation had been obtained in his experiment it would have been analysed further in an attempt to set up a reliable prediction system from first-generation response. Such a system might also predict the amount of fluctuation of response and the rate at which it falls off - this much was indicated by the negative correlation between first-generation response and regression coefficient. But once again a significant correlation would appear to depend on larger populations and greater replication than were present in the current experiment.

PART THREE: EFFECTS OF SELECTION INTENSITY

This final part of the thesis could be considered as an applied extension of the first two parts. The previous experiment had to be designed to enable the detection of extraneous effects such as natural selection and scale, and so to determine a satisfactory way of measuring the true genetic response. But this procedure meant that half of the labour required for the experiment must be put into Control lines and few commercial breeders have the desire or can afford to spend half of their costs on Controls. Hence the methods of analysing and predicting selection response that have been tested so far would have more applied value if they were re-tested under commercial conditions of selection, and the following experiment is designed to carry out this aim. A different type of Control population was used which did not require very much labour and the time which was thereby saved was spent on selecting at different intensities.

This particular parameter, selection intensity, was chosen for study because it has received little attention in the past. Just how little it has been studied by previous authors is shown by the smallness of literature review that follows.

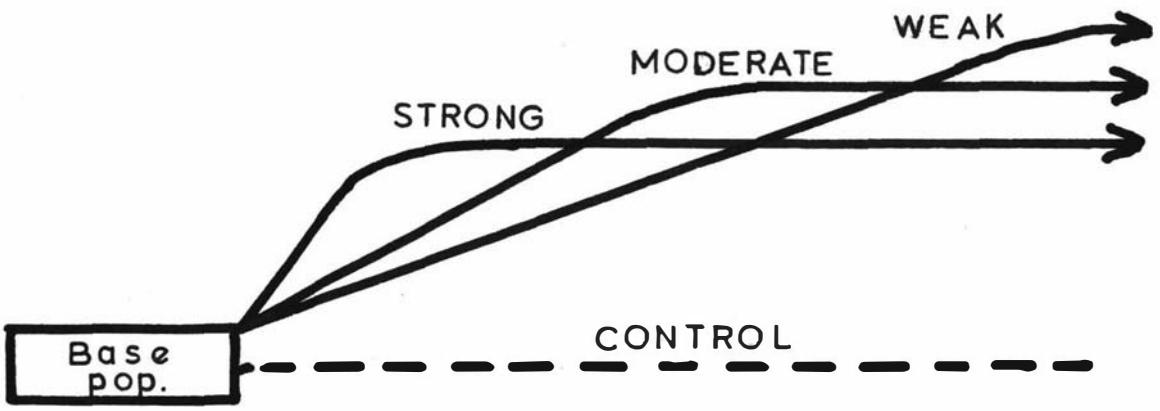
3.1 LITERATURE REVIEW:

The work of Clayton et al., (1957) is apparently the most detailed experiment done on effects of selection intensity. These authors selected for higher and lower chaetae number in Drosophila and at four selection intensities in each direction; there were three to five replications at each intensity. The results showed that over a short-term period the average rate of response was greatest when selection was most intense and lowest when selection was least intense. This is the result to be expected regardless of the type of gene action involved because high-intensity selection produces the largest selection differential

and therefore automatically it produces the largest selection response, at least over the first few generations. However, a subsidiary effect of high intensity selection is to quickly reduce genetic variability in the population and this in turn is expected to reduce the heritability more quickly than it would be reduced under low intensity selection. This leads to an expectation that the selection response from very intense selection will eventually lose its lead over that of a less intense system. Unfortunately this part of Clayton's experiment was stopped after five generations and no such reversal need have been expected in that short time. What may happen over a longer period is suggested by selection simulation on computers, as carried out by Martin and Cockerham (1960). These authors traced the selection response on a variety of genetic models which varied in number of loci, recombination index, type of gene action etc and of course, selection intensity. The response in the early generations of selection supported the results of Clayton, that intense selection produces the greatest response. As the programme continued however, this superiority was gradually reduced and sometimes eventually reversed. Thereafter the "heaviest" populations were those from the less intense selection system. The number of generations needed for the latter system to overtake the former, if it does so at all, apparently depends on the exact genetic models being used (Gill, 1965c; Young, 1966). The results of these simulation experiments therefore accord generally with genetic theory, although such simulations can never be claimed to represent the complex state in a real population. The present experiment undertook to attempt this necessary final step for assessing such theory and so it fills a large (and surprising) gap in the literature. But it is stressed that this aim of the experiment is no more important than the other aim given earlier - to apply the methods of analysing and predicting selection response which were used in the first experiment.

3.2 METHODS:

Organism: In this second experiment the Tribolium stock used was the original base population from Ruakura, and not the part-selected Heavy and Light ones. The reason for this change was that a large unselected base population should give a greater response to selection, for each intensity: the differences between the intensity responses should thus be measured more precisely.



Programme: Selection was made for heavier pupae at three separate intensities - Strong, Moderate, and Weak. It was continued for twenty generations, though this period was not to be broken up into equal units as was the first experiment. Instead the results would be analysed over periods of increasing length; namely over five, then over ten, then fifteen, and finally over twenty generations. It was considered that the results would have more applied value in this form.

There was no selection made for lighter pupae in this experiment but non-selection lines were extracted from the main ones at Generation 18, and carried

for two generations.

Control: It consisted of a population of 200 female and 200 male pupae being weighed each generation. A random fifty pairs of these were pair-mated, and eight pupae were taken from each progeny for weighing. This Control was started, not from the population used as base for the three intensity lines, but from the Heavy Ruakura base. This tactic was used simply to bring the Control weight closer to what the expected weights would be in the other three lines after a few generations of selection; it would therefore be more similar to them in matters of scale and natural selection. This explains why in Figure 10 the Control weight is higher than those of the selected lines for a very short time.

The Control population described above clearly differs in nature from that used in the earlier experiment. Because of the method of subdividing the total response and because of the commercial approach in this second experiment, this new type of Control was needed. It would not be as reliable as the previous one because it must cater for all three treatments - Strong, Moderate, and Weak. Hence it could not fulfil the dictum of Bray et al., (1962) to stay close to the selected lines in genotype - obviously it could not stay close to all three lines once they started to diverge. This is an admitted weakness of the Control, and it is common to most commercial Controls. Apart from this the Control complied with the requirements laid down for Tribolium in the experiment of Bray et al.

Replications: For each of the three treatments there were six replications. Thirty two female and 32 male pupae were randomly sampled and weighed in each replication every generation, exactly as in the previous experiment. No extra replications were carried in case of failures, nor were any subsequently needed.

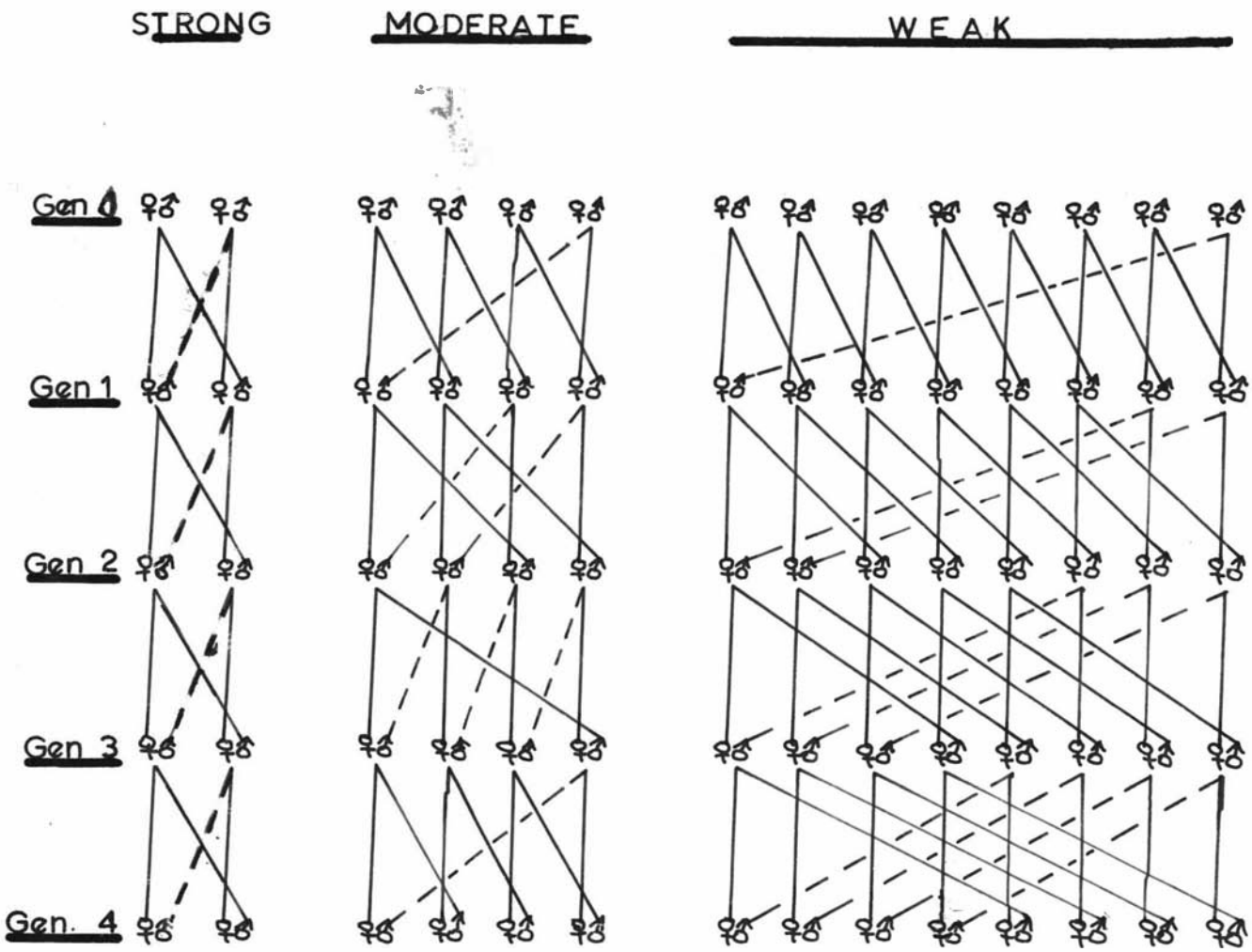
Selection procedure: (a) Eighteen samples, each of 32 female and 32 male pupae, were drawn from the base population. There were no significant weight differences among the eighteen, therefore they were randomly distributed among the three treatments, six replications to each.

(b) The populations were made equal in size in all three treatments because it was considered that commercial breeders were usually limited in the number of individuals they could handle. Hence selection intensity was varied by the number of parents chosen. In each replication of the Strong treatment, the two heaviest females and two heaviest males were selected out of the 32 pairs weighed; the four heaviest pairs were selected in each Moderate replication and the eight heaviest pairs in each Weak one.

(c) With so few parents being selected each generation it was vital that none of them die too soon - or at least without it being detected. If they did not all contribute an equal number of progeny to the next generation, the actual selection intensity may be far different from the theoretical one and this could seriously diverge the actual and theoretical responses. For this reason pair-matings were used throughout the experiment. There were two such matings in each Strong replication, four in each Moderate one and eight in each Weak one. Conversely, sixteen random female and sixteen random male pupae were weighed from each Strong pair-mating, eight pairs from a Moderate one, and four pairs from a Weak one. This gave a total of 32 pairs in each treatment per replication per generation. In all cases the heaviest female and male were selected from a pair-mating to be future parents.

(d) Inevitably the inbreeding rate would differ between the treatments, highest in the Strong, lowest in the Weak. This itself was not so important as the need to keep all inbreeding rates fairly constant within any one treatment. The allocation of partners for each new pair-mating was therefore an important point. It was desirable that the inbreeding rate be kept smooth between

generations within all replications, similar between replications within all treatments, and minimal for each treatment. The following circular mating plan was devised.



The expected inbreeding levels after twenty generations are approximately: Strong = 80%; Moderate = 45%; Weak = 24%. Kimura and Crow (1963) later devised a system that gives slightly slower inbreeding near the start but the difference between the two systems is very small at twenty generations.

Fig 10 : SELECTION RESPONSE

(FEMALE DATA)

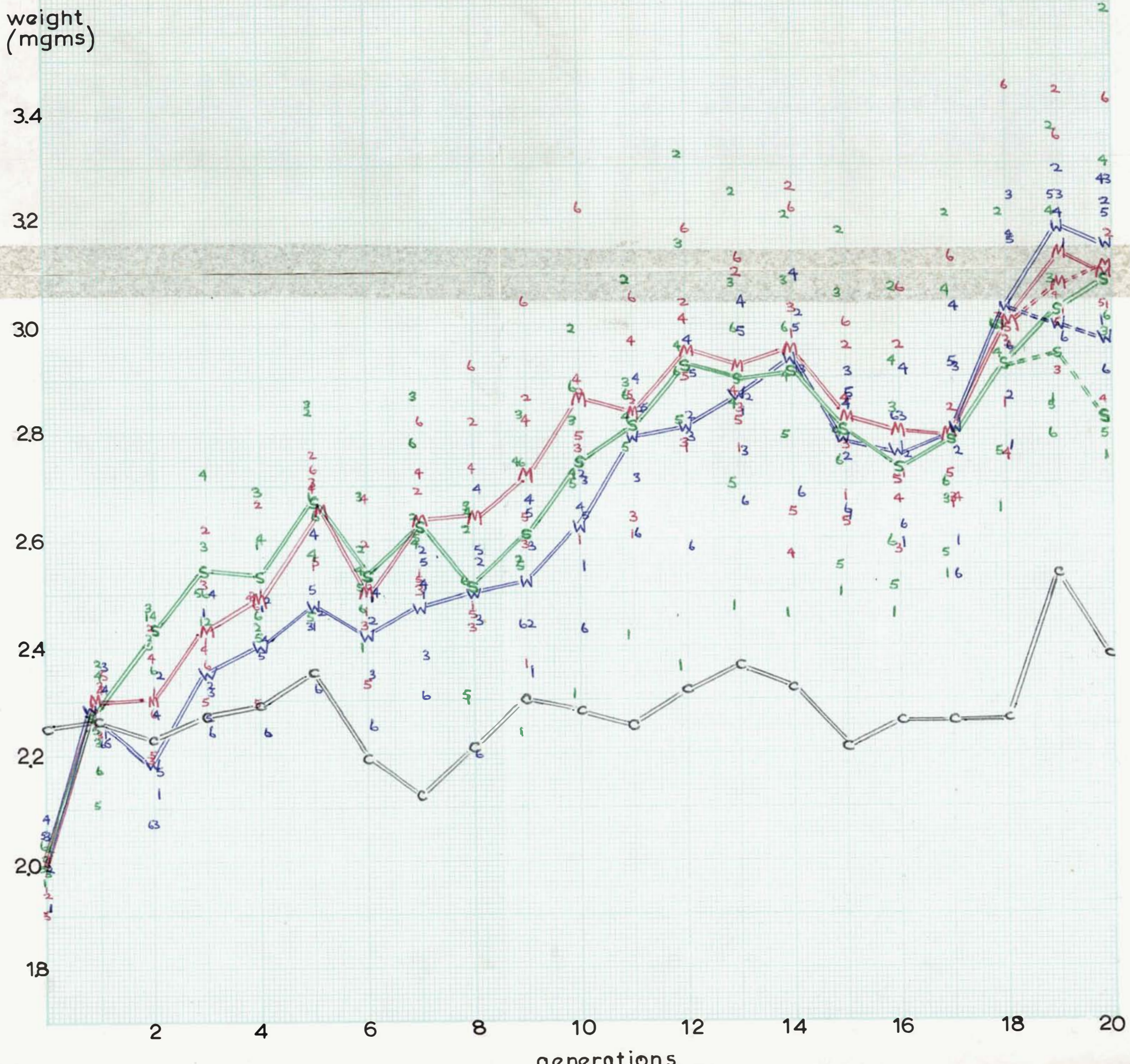
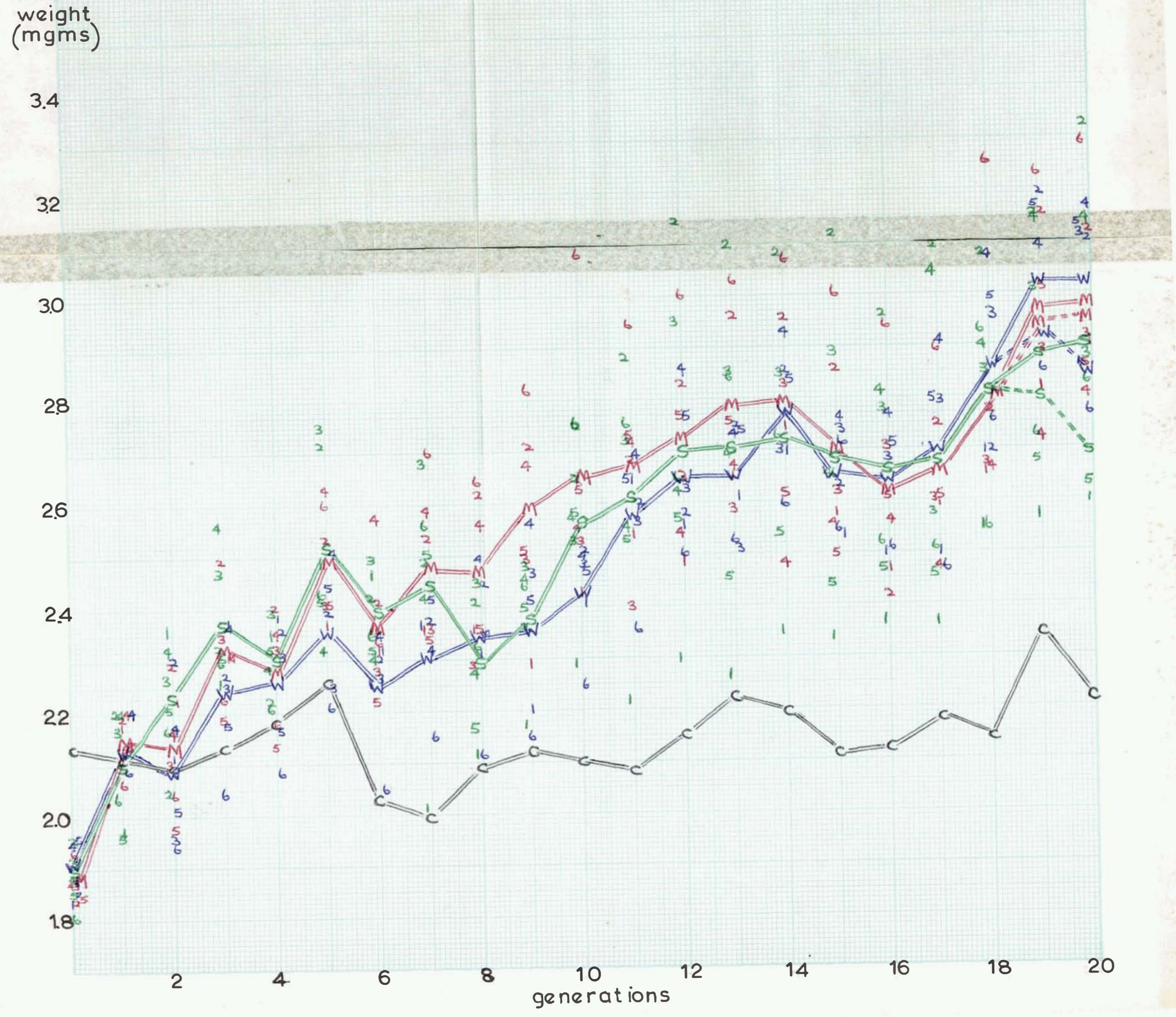


Fig 10 : continued

(MALE DATA)



3.3 RESULTS:

This second experiment was run about simultaneously with the first one, beginning in September 1963 and ending in June 1965. It also suffered the same ordeal by overheating in about the first generation and so was restarted from reserve matings; but apart from this the twenty generations of selection were completed with little trouble. As the experiment neared its end a problem of low fertility developed in the Strong treatment, possibly due to the high inbreeding rate. This problem did not cut down the number of pupae weighed but it did mean waiting longer before the quota of weighings was reached. Automatically the generation interval was increased for the Strong treatment and since the cycles of Strong, Moderate and Weak all had to be kept in step with the Control, those of Moderate and Weak had to be allowed to increase also. The number of days required for each generation (averaged over all three treatments) is given in Table 23, and is worked out on the same basis as for the first experiment.

Table 23: Generation Interval (in days)

Generation	Interval
0 to 1	30
1 to 2	33
2 to 3	32
3 to 4	30
4 to 5	33
5 to 6	31
6 to 7	31
7 to 8	31
8 to 9	31
9 to 10	30
10 to 11	31
11 to 12	30
12 to 13	31
13 to 14	30
14 to 15	31
15 to 16	34
16 to 17	34
17 to 18	32
18 to 19	33
19 to 20	32
mean	31.5

3.31 Analysing selection response

(A) PRESENTATION OF BASIC DATA:

Mean weight (Figure 10 and Table 24): All three treatments give a fairly marked response till about Generation 14, notwithstanding the large fluctuations. Then all three lose weight for a short period before resuming a steady response. The Control population shows considerable fluctuation also but no environmental trends; when this Control variation is taken into account (Fig. 12) some of the treatment fluctuations are removed but others are not. For example the depression after Generation 14 is converted to a plateau. The reason for this plateau can only be surmised but the fact that it affects all three treatments suggests that an unnoticed change of diet or climate may be responsible, though such explanation must assume that the same change has not affected the Control population to an equal degree.

At first Strong gives the fastest response followed by Moderate, but after seven generations Moderate takes the lead. By Generation 17 the Weak populations are heaviest and thus the original order is reversed. This is the pattern of response that was predicted by many of the genetic models used in the simulation experiments of Martin and Cockerham (1960), Gill (1965c) and Young (1966). Latent heritability would be the same for all three treatments at the start of selection and therefore Strong, with the highest selection differential and most rapid increase of important genes, should give the greatest response. However, its genetic variation would also automatically fall at the fastest rate because so few parents were being selected each generation, therefore the heritability would fall off faster here than in the other treatments and with it the rate of response. Conversely Weak has the smallest selection differential and so would give the smallest response in the early generations. But since its inbreeding

rate is lower it loses fewer favourable genes by genetic drift and so the heritability and rate of response would be maintained for much longer. Indeed there is little suggestion from the graphs that they have fallen very much at all over the whole twenty generations.

The two generations of non-selection at the end of the experiment fluctuate too greatly to show any clear pattern of response. Certainly they do not keep up with their respective selection progenitors but conversely they do not tend to regress very much. It cannot be concluded therefore that there is evidence of natural selection operating or that the response to artificial selection is based on epistatic variance (Griffing, 1960). These non-selection lines are not used again during the analysis of results.

Phenotypic variance (Figure 11 and Table 25): As was the case for the first experiment, data for the graph was drawn from that of Table 25 by merely averaging the twelve values (6 replications X 2 sexes) in each generation. Though such twelve values may be heterogeneous (and consequently their mean may not have too much usefulness) it was considered that the pattern of these means over several generations would at least show any trends of increasing or decreasing variance.

In Figure 11 there is a slight rise of value in all three treatments till about the ninth generation but thereafter no clear pattern in any of them. The slight rise may be the result of scale effects but if this is so it is being counteracted in later generations, perhaps by a declining genetic variation. There is insufficient scale evidence that variability would be better measured in terms of the coefficient of variation and so the data have not been converted.

Realised heritability (Figure 11 and Table 26): This parameter was once again calculated as the ratio of selection response to selection differential. The numerators were adjusted to cater for variation of the Control population,

Fig 11 : RESPONSE IN OTHER PARAMETERS

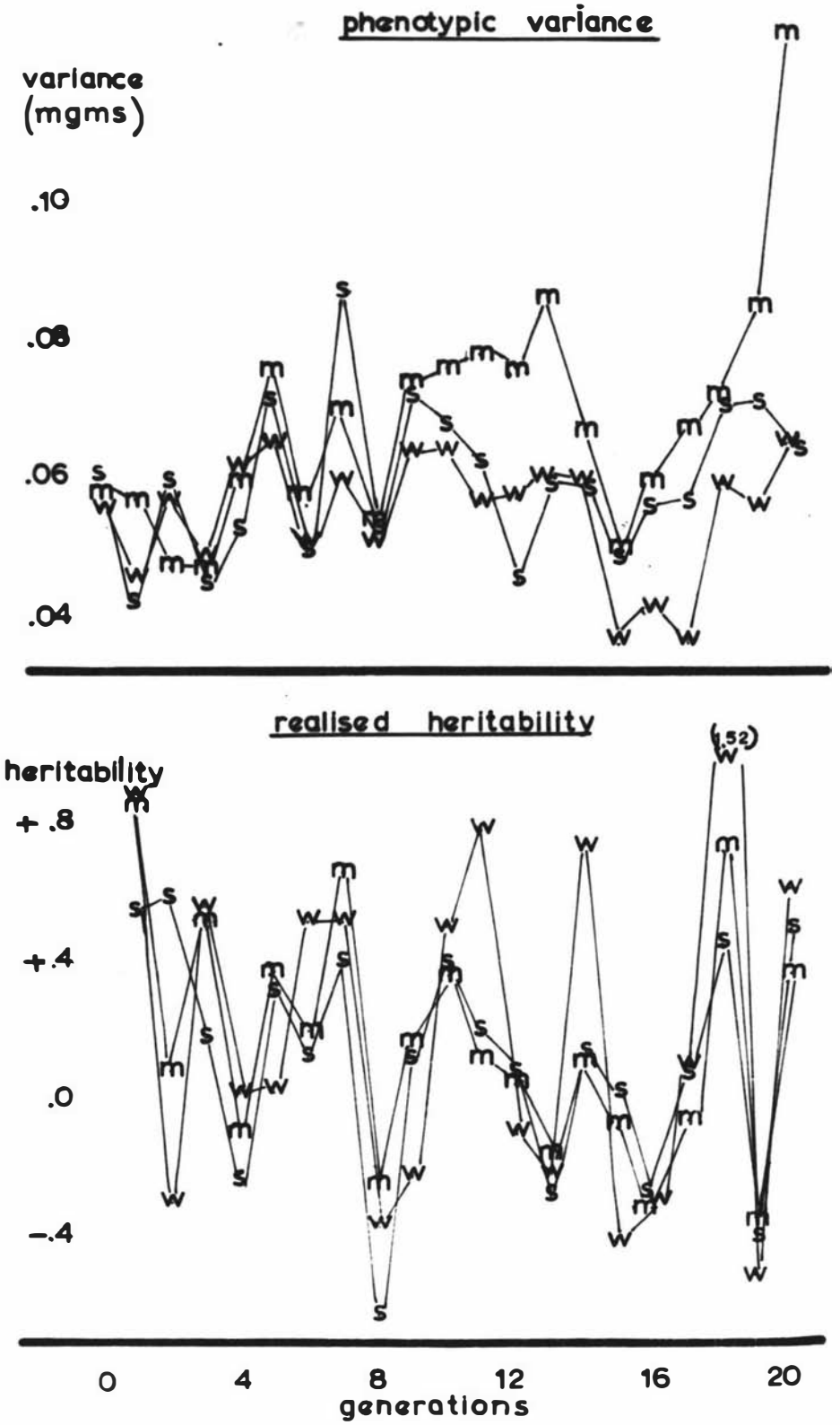


Table 25: Phenotypic variances a) Female data

	STRONG							MODERATE							WEAK						
Rep	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean
Gen																					
0	.062	.054	.062	.058	.055	.048	.056	.054	.039	.051	.058	.053	.084	.056	.066	.034	.040	.087	.060	.033	.053
1	.063	.032	.034	.050	.040	.029	.041	.049	.063	.057	.043	.039	.064	.052	.046	.038	.037	.035	.066	.042	.044
2	.052	.069	.051	.045	.065	.057	.056	.044	.054	.038	.036	.057	.042	.045	.066	.049	.095	.061	.068	.034	.062
3	.052	.044	.030	.035	.042	.035	.040	.037	.049	.072	.047	.041	.042	.048	.021	.040	.044	.059	.066	.068	.050
4	.067	.063	.059	.030	.040	.073	.055	.033	.088	.066	.060	.032	.089	.061	.049	.123	.063	.047	.064	.072	.070
5	.082	.052	.078	.111	.037	.050	.068	.070	.075	.052	.065	.083	.088	.072	.094	.055	.097	.090	.035	.063	.073
6	.055	.100	.048	.083	.037	.030	.059	.088	.053	.079	.033	.069	.037	.060	.072	.075	.035	.042	.071	.037	.055
7	.152	.068	.067	.062	.103	.068	.087	.092	.043	.047	.074	.072	.094	.070	.027	.063	.057	.066	.079	.083	.062
8	.061	.046	.052	.095	.115	.040	.068	.035	.037	.042	.089	.039	.100	.057	.060	.043	.043	.055	.066	.052	.053
9	.048	.131	.096	.073	.081	.052	.080	.086	.054	.040	.082	.091	.089	.074	.083	.081	.043	.061	.045	.090	.067
10	.044	.124	.101	.086	.067	.074	.083	.040	.093	.074	.090	.106	.052	.076	.041	.050	.060	.088	.075	.055	.062
11	.046	.096	.111	.046	.035	.036	.062	.090	.099	.098	.056	.058	.079	.080	.037	.045	.043	.059	.053	.064	.050
12	.057	.036	.037	.043	.018	.051	.040	.107	.101	.082	.045	.055	.080	.078	.055	.065	.064	.059	.055	.048	.058
13	.105	.048	.057	.032	.071	.048	.060	.142	.048	.066	.093	.118	.074	.090	.086	.040	.072	.050	.046	.057	.059
14	.056	.085	.050	.045	.065	.029	.055	.065	.048	.039	.059	.101	.091	.067	.025	.065	.063	.034	.072	.047	.051
15	.051	.024	.077	.056	.042	.030	.047	.081	.047	.038	.034	.060	.052	.052	.042	.033	.036	.053	.038	.019	.037
16	.022	.063	.111	.059	.057	.050	.060	.073	.074	.075	.043	.061	.084	.069	.035	.044	.046	.029	.041	.026	.037
17	.029	.063	.054	.097	.040	.067	.058	.065	.024	.055	.069	.100	.131	.074	.040	.038	.037	.048	.065	.028	.043
18	.056	.090	.107	.070	.057	.096	.079	.041	.062	.134	.104	.100	.030	.078	.032	.047	.053	.108	.075	.133	.075
19	.039	.114	.089	.062	.044	.109	.076	.094	.041	.149	.072	.152	.040	.091	.053	.105	.059	.083	.046	.032	.063
20	.070	.076	.102	.045	.058	.030	.063	.117	.154	.084	.181	.151	.069	.126	.101	.056	.075	.036	.082	.073	.071

Table 25: (Continued) b) Male data

	STRONG							MODERATE							WEAK						
Rep	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean
Gen																					
0	.057	.068	.058	.066	.039	.087	.063	.073	.077	.049	.043	.055	.054	.059	.052	.053	.069	.056	.047	.073	.058
1	.059	.042	.035	.033	.037	.051	.043	.039	.058	.046	.079	.068	.082	.062	.047	.059	.040	.048	.040	.053	.048
2	.068	.067	.051	.045	.106	.042	.063	.053	.067	.042	.037	.032	.072	.050	.050	.050	.070	.069	.037	.043	.053
3	.051	.068	.040	.034	.023	.073	.048	.029	.049	.066	.043	.031	.049	.044	.032	.048	.053	.077	.042	.040	.048
4	.066	.051	.058	.037	.045	.054	.052	.041	.073	.050	.068	.046	.070	.058	.045	.059	.079	.034	.046	.069	.055
5	.072	.064	.057	.108	.083	.062	.074	.087	.093	.094	.042	.052	.098	.078	.063	.040	.060	.049	.069	.060	.057
6	.037	.039	.055	.047	.030	.030	.040	.069	.041	.057	.033	.058	.055	.052	.054	.027	.033	.040	.070	.046	.045
7	.133	.075	.062	.073	.133	.055	.088	.054	.057	.045	.060	.065	.134	.069	.035	.069	.068	.043	.064	.070	.058
8	.023	.020	.034	.069	.052	.028	.038	.025	.052	.025	.049	.035	.108	.049	.031	.053	.033	.057	.077	.047	.050
9	.026	.096	.040	.063	.072	.103	.067	.046	.095	.047	.100	.109	.085	.073	.061	.068	.063	.047	.092	.037	.061
10	.026	.079	.055	.040	.069	.057	.054	.026	.057	.089	.083	.135	.063	.075	.050	.058	.069	.091	.069	.048	.064
11	.032	.112	.080	.066	.059	.031	.063	.064	.079	.056	.077	.135	.046	.076	.042	.062	.082	.069	.059	.059	.062
12	.040	.021	.063	.062	.069	.048	.051	.079	.043	.094	.067	.077	.075	.073	.048	.085	.055	.047	.074	.033	.057
13	.045	.050	.101	.040	.088	.044	.061	.110	.057	.095	.091	.095	.049	.083	.059	.033	.092	.054	.081	.050	.061
14	.062	.056	.075	.048	.067	.078	.064	.117	.090	.056	.038	.054	.047	.067	.039	.064	.083	.093	.068	.061	.068
15	.051	.037	.102	.049	.027	.029	.049	.055	.023	.037	.053	.057	.051	.046	.055	.028	.032	.048	.039	.017	.036
16	.051	.039	.061	.068	.050	.041	.052	.052	.042	.031	.035	.081	.059	.050	.024	.074	.047	.046	.034	.047	.045
17	.050	.079	.060	.054	.043	.056	.057	.032	.036	.049	.053	.078	.126	.062	.037	.029	.028	.032	.024	.027	.030
18	.078	.070	.076	.070	.041	.063	.066	.025	.040	.100	.085	.093	.052	.066	.040	.065	.059	.104	.054	.060	.064
19	.073	.086	.061	.043	.058	.086	.068	.104	.055	.114	.080	.079	.051	.081	.055	.035	.075	.038	.021	.065	.048
20	.045	.103	.073	.076	.063	.051	.068	.088	.142	.064	.133	.226	.085	.123	.072	.064	.072	.039	.080	.036	.060

according to the formula of Kojima and Kelleher (1963).

$$\Delta\bar{Y}_i = (\bar{Y}_i - \bar{C}_i) - (\bar{Y}_{i-1} - \bar{C}_{i-1})$$

where $\Delta\bar{Y}_i$ = mean response at Generation i

$\bar{Y}_i$ = mean weight at Generation i in selection line

$\bar{C}_i$ = mean weight at Generation i in Control line.

In Figure 11 all three treatments show big fluctuations, despite the fact that each point on the graph is once again the mean for male and female data in all six replications. A survey of the data suggested that this fluctuation arose mainly from the denominator of the ratio, that is from the selection differential. It is therefore probably related to the fluctuating phenotypic variability of the population and like the latter it is probably caused by sampling error rather than by any erratic genetic behaviour. The fluctuations are so big in the Weak treatment that they mask any trends of change in the heritability value but in both Moderate and Strong there appears to be a decrease of value as selection proceeds. This general conclusion would be about the limit of information that could be taken from these graphs and such conclusions should be quantified by a statistical analysis of the data, exactly as was required in the first experiment.

Table 26: Realised heritabilities a) Female data

	STRONG							MODERATE							WEAK						
Rep	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean
Gen																					
1	.84	.72	.52	.75	.29	.42	.59	.81	1.23	.61	.89	1.16	.65	.89	.86	1.29	1.11	.73	.55	1.04	.93
2	.55	.50	1.07	.46	1.17	.82	.76	.38	.52	-.04	.50	-.48	.18	.18	-.39	.80	-1.79	.00	-.09	-.86	-.39
3	-.18	-.05	.19	.53	.10	.23	.14	.13	.43	1.33	-.15	.23	.18	.36	1.11	-.50	.87	.81	.18	.67	.52
4	.40	-.16	.26	-.50	-.40	-.23	-.10	.27	.11	-.14	.26	-.11	.36	.12	-.11	1.00	.47	-.42	.39	-.12	.20
5	.06	1.30	.27	-.36	-.05	.32	.26	-.04	.07	.44	.50	.81	.49	.38	-.45	-.50	-.12	.65	.20	.09	-.02
6	-.32	-.32	-.02	.39	.48	-.02	.03	.30	-.03	-.46	.61	-.19	-.18	.01	.66	.61	.20	.17	.65	.41	.45
7	-.23	.28	.81	.33	1.14	1.36	.61	.74	.47	.61	.60	1.00	1.23	.77	.35	.80	.65	.42	.65	.68	.59
8	-.09	-.30	-.79	-.08	-1.52	-.95	-.62	-.44	.14	-.41	-.29	-.43	.04	-.23	-.47	-.61	-.04	.36	-.29	-1.27	-.39
9	-.30	-.28	.20	.00	.31	.46	.06	-.68	-.17	.18	.00	.37	.09	-.03	-1.00	-1.17	.25	-.55	-.09	.70	-.31
10	.20	.76	.02	.00	.49	.44	.32	.76	.08	.95	.32	.37	.56	.51	1.00	1.15	1.17	.00	.06	.04	.57
11	.48	.20	.19	.30	.33	.07	.26	.20	.04	-.33	.39	.30	-.56	.01	1.36	.75	.20	.76	1.00	1.00	.84
12	-.28	.43	.43	.35	-.04	-.09	.13	.32	.41	.29	-.08	-.04	.23	.19	-.42	-.47	.05	.00	.11	-.39	-.19
13	.12	-.36	-.48	-.55	-.74	.09	-.32	-.13	.04	.03	-.73	-.67	-.30	-.29	-.06	.00	-.41	.12	.11	.24	.00
14	.05	.00	.12	.33	.37	.15	.17	.62	.80	.88	-1.04	-1.00	.47	.12	.92	1.12	.95	.53	.19	.21	.65
15	.38	.21	.28	.00	-.50	-.67	-.05	-.56	-.65	-.82	1.60	.18	-.37	-.10	-1.13	-1.33	.45	-1.75	-.04	.44	-.56
16	-.22	-.59	-.69	.32	-.45	-1.25	-.48	-.03	-.14	-1.04	-1.00	.09	.40	-.29	-.56	-.27	-.72	.15	-.69	-.67	-.46
17	.29	.22	.33	.36	.15	.31	.28	-.17	-.33	.23	.00	.04	-.11	-.06	.00	.00	.53	1.22	.56	-.69	.27
18	.54	.00	.74	-.18	.68	.83	.43	.49	.65	1.07	.29	.96	.97	.74	1.06	.53	2.13	1.17	.96	3.58	1.57
19	-.20	-.28	-.32	.00	-.65	-1.52	-.49	-.33	.38	-.82	.07	-.82	-1.94	-.58	.53	.84	-1.35	-.82	-1.27	-2.50	-.76
20	.20	.86	.40	.83	.32	.90	.59	.35	-.41	.82	-.23	.44	.81	.30	.04	.32	.82	.91	.54	.64	.55

Table 26: (Continued) b) Male data

	STRONG							MODERATE							WEAK						
Rep	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean	1	2	3	4	5	6	mean
Gen																					
1	.25	.55	.64	.83	.31	.46	.51	.56	.79	.79	1.20	1.10	.47	.82	1.19	.93	.78	1.04	.76	.69	.90
2	1.08	-.58	.32	.45	.68	.37	.39	.26	.37	-.14	-.12	-.47	-.03	-.02	.09	.87	-.89	-.09	-.69	-.62	-.22
3	-.36	.51	.48	.53	.10	.24	.25	.18	.57	1.00	.65	1.00	.48	.65	1.23	-.41	1.25	.56	.60	.35	.60
4	.10	-.39	-.39	-1.14	-.17	-.35	-.39	-.40	-.50	-.19	-.07	-.38	-.31	-.31	-.33	.23	.10	-.58	-.35	-.06	-.16
5	.10	1.27	.61	-.09	.05	.32	.38	.00	.17	.04	.63	.58	.77	.36	-.90	-.20	-.56	.75	1.00	.29	.06
6	.50	-.15	-.07	.64	.30	.25	.24	1.20	.21	.24	.64	.14	-.11	.39	.96	.67	1.00	.40	.18	.29	.58
7	-2.16	.61	.64	.46	.96	.72	.20	.37	.50	.37	.22	.49	1.28	.54	.40	.55	-.06	.09	.95	.70	.44
8	.02	-.35	-.79	-.47	-1.23	-1.17	-.66	-.29	-.06	-.55	-.35	-.19	-.32	-.29	-.83	-.09	.10	.35	-.82	-1.08	-.39
9	.09	-.29	.00	.39	.47	.21	.14	-.42	.17	.74	.26	.57	.54	.31	-1.00	-.57	.50	.14	.20	.00	-.12
10	.42	.70	.22	.35	.50	.57	.46	.61	-.16	.19	-.67	.45	.85	.21	1.09	.67	.11	-.23	.28	.81	.45
11	-.24	.50	.40	.00	-.07	.14	.12	.68	.23	-.28	.59	.37	-.63	.16	1.39	.62	.52	.68	.70	.50	.73
12	.03	.40	.39	.00	-.08	-.72	.00	-.46	.22	.47	-.71	-.12	-.04	-.11	-.88	-.42	.00	.61	.22	.30	-.03
13	-.25	-.48	-.61	.02	-.55	.36	-.25	.24	.27	-.48	.18	-.29	-.12	-.03	-.09	.37	-1.19	-1.25	-.45	.16	-.46
14	.33	.00	.09	.30	.22	-.26	.11	.41	.12	.51	-.37	-.32	.26	.10	.57	.76	1.05	1.77	.46	.39	.83
15	.16	.29	.27	-.10	-.06	.00	.09	-.31	-.08	-.48	.56	-.09	.04	-.06	-.47	-.93	.50	-.82	-.26	.23	-.29
16	.06	-.61	-.33	.60	.09	-.56	-.12	-.33	-1.00	-.75	.00	.32	-.32	-.35	-.21	.06	-.41	.00	.00	-.36	-.15
17	-.16	.16	-.96	.71	-.18	-.18	-.10	.18	-.07	.60	-.58	-.20	-.31	-.06	-.60	-.17	.28	.36	.20	-.53	-.08
18	.60	.12	.60	-.28	.36	1.35	.46	.42	.35	.44	.80	1.15	.98	.69	1.53	.40	1.40	1.33	1.92	2.27	1.47
19	-.49	-.35	-.10	.11	-.23	-.85	-.32	-.21	.71	.00	-.39	.00	-.88	-.13	.00	1.50	-.87	-.79	-.25	-1.30	-.28
20	.37	.62	.00	.42	.30	.72	.40	.47	.37	.38	.70	-.03	.62	.42	.43	.31	1.22	1.11	.77	.27	.68

(B) METHODS OF ANALYSING DATA:

The same approach was used as in the first experiment, that is to divide the total programme into periods and then to measure the mean response and fluctuations and trends during each period. Because they showed much similarity the female and male data in each replication were pooled in every generation, but the six replications themselves were analysed separately.

Average rate of response: In the first experiment two methods of measuring the mean response were compared. The first method involved fitting a straight regression line to the generation mean weights and using the value of regression coefficient as a measure of the average selection response per generation. This method was criticised on the grounds that there was a considerable negative correlation between the deviations of adjacent points from the regression line. The second method studied was to measure average response per generation simply as the total response divided by number of generations. This method can be criticised on the grounds that it involves only the mean weights of the first and last generations, so if either of these is estimated wrongly the error will have a very big effect on the calculation. However, both sets of results were presented because there was considerable interest in comparing the degree of difference between them.

In this second experiment the two methods were again compared for suitability. The test carried out below is a test for non-linearity of the regression lines applied to each replication. The test was applied only to the total twenty generations of response and not to the periods of five, ten, or fifteen generations, as it was considered that the former would be most likely to show non-linearity if any existed. The generation means used for analysis were all corrected for variation of the Control population (Fig. 12) before the regression lines were fitted. All three treatments were so analysed for non-linearity (Table 27).

Fig 12: SELECTION RESPONSE ADJUSTED FOR CONTROL

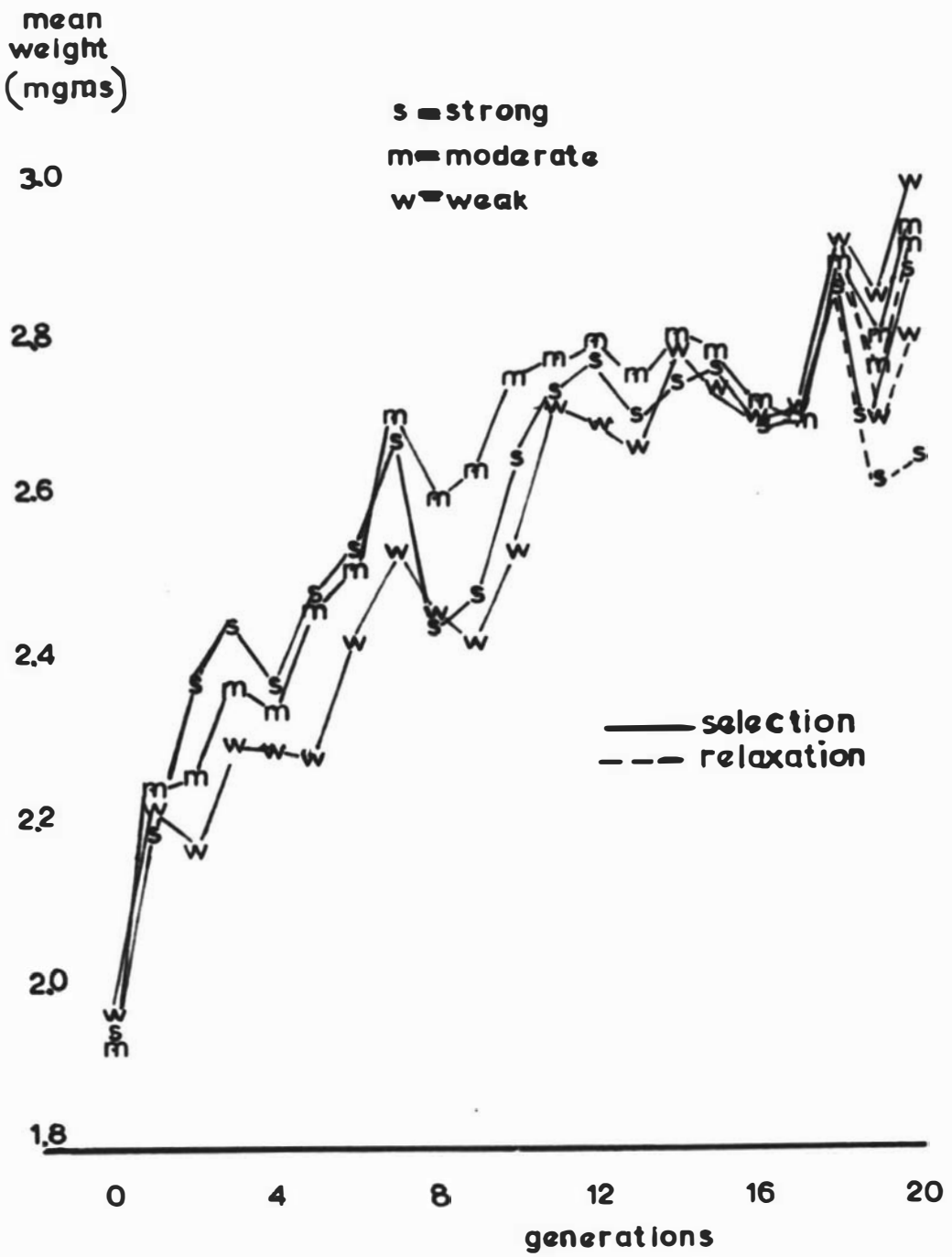


Table 27: Deviations from linear regression

Strong

Source of variance	df	MS	F	P
Replications	5	0.4900		
Regression	1	4.8500		
Deviation from Regr.	19	0.0832	4.06	.001
Residual	100	0.0205		
Total	125			

Moderate

	df	MS	F	P
Replications	5	0.2920		
Regression	1	6.1600		
Deviation from Regr.	19	0.0879	6.37	.001
Residual	100	0.0138		
Total	125			

Weak

	df	MS	F	P
Replications	5	0.1880		
Regression	1	8.0600		
Deviation from Regr.	19	0.0363	5.85	.001
Residual	100	0.0062		
Total	125			

In all three treatments the mean square for Deviation is highly significant compared with that for Residual. This indicates that there is a highly significant non-linear effect, which in turn weakens the usefulness of linear regression as a measure of average response. Consequently the author has measured average response throughout the present experiment only in terms of the total response

divided by number of generations. The degree by which the two methods differ in their estimates may be judged from Table 28 below, which deals with the whole twenty generations. Regardless of this degree of difference however, the regression method was not used because definite curvilinearity (see later analysis) contributed to the significant deviations above. Were this not so, the regression method would be a suitable measure of average response in the least squares sense, even despite the significant non-linearity.

Table 28: Two methods of measuring mean selection response (Generations 0 to 20, in mgms per gen)

	Rep 1	2	3	4	5	6
STRONG						
regression method	.0130	.0571	.0315	.0394	.0219	.0314
total/20	.0322	.0692	.0555	.0587	.0347	.0450
difference	-.0192	-.0121	-.0240	-.0193	-.0128	-.0136
difference/2nd est.	60%	17%	43%	33%	37%	30%
MODERATE						
regression method	.0299	.0395	.0317	.0220	.0379	.0580
total/20	.0477	.0577	.0477	.0392	.0487	.0645
difference	-.0178	-.0182	-.0160	-.0172	-.0108	-.0065
difference/2nd est.	37%	32%	34%	44%	22%	10%
WEAK						
regression method	.0318	.0398	.0475	.0460	.0486	.0371
total/20	.0472	.0572	.0562	.0550	.0532	.0402
difference	-.0154	-.0174	-.0087	-.0090	-.0046	-.0031
difference/2nd est.	33%	30%	15%	16%	9%	8%

On every occasion the value of average response is less when measured from linear regression than it is when measured from simply dividing the total response by number of generations. It is very likely that the reason for this consistently smaller value lies in the plateauing of response between about Generations 12 and 17, which affected all three treatments. Such plateauing would tend to reduce

the value of regression coefficient but would not influence the value of average response as measured by the second method. Similarly the observation that Weak treatment had smallest mean weight before the plateau and largest mean weight after it, indicates that Weak was least affected by this plateau; and this explains why the percentage difference between the two methods of estimating average response is fairly consistently least in Weak.

The two methods differ in their estimates of average response by fairly large percentages, and only the second method (that is non-regression) has been used during the remainder of the present experiment. It is noted in passing however, that the two methods gave virtually the same rankings of treatment responses.

Fluctuations of response: The method of measuring fluctuations followed almost automatically from the decision to measure mean response directly from the total. The response at each generation was measured from

$$\Delta\bar{Y}_i = (\bar{Y}_i - \bar{C}_i) - (\bar{Y}_{i-1} - \bar{C}_{i-1})$$

The standard deviation of these separate responses about the period mean response was then calculated.

Trends of response: Indications as to whether the selection response was increasing or decreasing as generations pass by were again obtained by fitting a regression line to the separate responses as calculated by the formula above. If these responses were all equal then the coefficient of regression would be zero, but if they tended to increase or decrease the coefficient would be respectively positive or negative.

Finally it is repeated that to give the experiment more applied value, the

results were analysed separately over the first five generations of selection, the first ten, the first fifteen, and finally over all twenty generations. Most of the interest lies in comparing Strong, Moderate and Weak responses within each period in turn, though some information on trends of each treatment in time can be gained by comparing different periods.

All parameters were again calculated separately for each replication.

(C) THE SELECTION RESPONSE:

Response mean: The mean responses per generation for each period are given in Table 29. The upper half of this table contains results which have not been adjusted for Control performance, and the lower half contains the adjusted results. The probability that there is no real difference between any two treatments (based on replication variance) is given in brackets between them in this lower half but of course the same probabilities apply exactly whether there is a Control or not.

Table 29: Mean response in each period (mgms per generation)

Generations	0 to 5	0 to 10	0 to 15	0 to 20
	a) Response unadjusted for Control			
Selection intensity				
Strong	.132	.071	.054	.054
Moderate	.131	.083	.055	.055
Weak	.090	.057	.051	.056
Control	+.023	.000	-.002	+.002
	b) Adjusted response			
Strong	.109 ± .012 (.9)	.071 ± .007 (.3)	.056 ± .007 (.9)	.052 ± .006 (.8)
Moderate	.108 ± .007 (.01)	.083 ± .008 (.02)	.057 ± .004 (.4)	.053 ± .004 (.9)
Weak	.067 ± .007 (.02)	.057 ± .004 (.2)	.053 ± .002 (.7)	.054 ± .003 (.8)

(Note: Third row of brackets relates to Strong-Weak differences)

The Control gives a large adjustment only to the first-period results. It contains no trends of change and its fluctuations that show so clearly on the graph are almost wholly cancelled out over the longer periods.

The trend within each treatment becomes fairly clear from this table and it

accords quite well with additive genetic theory. In Strong and Moderate the average response over the first five generations is considerably greater than that over the longer periods. The differences between periods are not so great from then onwards but nevertheless the average response consistently becomes less as the length of period under consideration becomes greater; in other words the response falls off with time. In the Weak treatment the first five generations again give the biggest response but this response does not fall away so markedly as in the other treatments.

Within any one period there are few significant differences between the average responses of the three treatments. The most noticeable difference is that Weak gives a significantly poorer response over about the first ten generations. Other differences might of course have proved to be statistically significant if more animals had been weighed and more replications had been carried: tentatively it can be said that there is a trend for the order of superiority to be

Strong then Moderate then Weak over 5 generations
Moderate then Strong then Weak over 10 and 15 generations
Weak then Moderate then Strong over 20 generations.

The value of these results would be greater if they were more precise but some conclusions can be given. If mean response was the only criterion to consider for a commercial breeding programme the results suggest that if selection is to be continued for only a short time then an intense treatment is the best one to use; but if this period of selection is to be longer the intensity should conversely be reduced. The best overall advice to be taken from the above results however, probably would be to carry out Moderate selection regardless of the period involved. Moderate is never more than very slightly below the best treatment and over the middle distances is itself the best by fairly significant amounts.

The genetical interpretation of these results can be checked by calculating

the realised heritabilities. The results are given in Table 30 and as with all results from now on, adjustment for variation of the Control population was made before the calculations were carried out.

Table 30: Realised heritability in each period

Generations	0 to 5	0 to 10	0 to 15	0 to 20
Strong	.276 ± .039 (.2)	.178 ± .027 (.02)	.128 ± .019 (.05)	.112 ± .017 (.05)
Moderate	.343 ± .029 (.1)	.280 ± .020 (.1)	.186 ± .010 (.7)	.157 ± .004 (.01)
Weak	.242 ± .039 (.6)	.214 ± .021 (.3)	.194 ± .012 (.02)	.216 ± .017 (.01)

This table shows that over the first five generations there are no significant differences among the treatments, though it must be admitted that they do approach significance. The result which is most surprising is that Weak should be so much lower than Moderate over this short period and no simple genetical reason is obvious to explain it. However, it is as likely to be merely a weakness of the experimental set-up as it is to be a genetical anomaly.

Over the three longer periods of selection Strong heritability is significantly less than that of Moderate, and significantly less than that of Weak over fifteen and twenty generations. The only other significant difference is between Moderate and Weak over the whole twenty generations. These results accord very well with those in the previous table and with genetic theory - that the rate of response is reduced in all three treatments with time and most rapidly so at the start, and that this response is also reduced most rapidly at the highest rate of

inbreeding, that is in the Strong treatment. This latter pattern is probably accelerated by inbreeding depression because characters that are related to an individual's fitness tend to suffer with inbreeding, and pupal weight is likely to be such a character. It is impossible to separate this effect of inbreeding from that due merely to a decreasing genetic variation, without carrying extra populations subjected to inbreeding but not to selection.

Standard deviation: Standard deviation was used as a measure of fluctuation of separate generations about their mean. It was calculated separately for each replication in turn (Fig. 14) and then all six values were averaged. The validity of averaging such values was first checked by Bartlett's test of homogeneity on the variances from which these standard deviations were derived. This test was applied only to the variances calculated for the twenty-generation period and not to those variances calculated for the shorter periods: the results are given in Table 31.

Table 31: Test for homogeneity of variances
(E = value of $\frac{1}{C}\left\{f \log_{10} s^2 - \Sigma f_i \log_{10} s_i^2\right\}$
P = probability that E is not significant, based on χ^2, 5 d.f.)

	E	P
Generations 0 to 20		
Strong	4.63	.5
Moderate	3.86	.6
Weak	3.14	.7

There is no evidence of heterogeneity in any of the three treatments and so it is meaningful to average their standard deviations, as is done in Table 32.

Table 32: Trends and fluctuations of response

Generations	0 to 5	0 to 10	0 to 15	0 to 20
		Standard deviation		
Strong	.176	.194	.170	.173
Moderate	.156	.149	.150	.155
Weak	.156	.139	.134	.140
		Regression coefficient		
Strong	-.0512	-.0164	-.0094	-.0050
Moderate	-.0418	-.0126	-.0108	-.0056
Weak	-.0473	-.0110	-.0055	-.0023

The value of standard deviation is consistently greatest in Strong and smallest in Weak. Not too much importance should be placed on the consistency of this trend from one period to the next because these periods all contain some generations in common, and furthermore there is a confounding effect between decrease in genetic variance and decrease in buffering ability. The pattern of decrease from Strong to Weak could obviously have arisen quite by chance in the first few generations and if this pattern was sufficiently powerful it would persist even when the longer periods were being analysed. For this reason only the longest period (twenty generations) has been investigated to determine if the pattern was significant. This significance was measured by an F-test comparison of variances, the variance values used being taken as s^2 from the homogeneity expression above. This s^2 value is the pooled variance over all six replications and consequently the pooled degrees of freedom are also used. The variances compared were

Strong $s_S^2 = .03085$

Moderate $s_M^2 = .02447$

Weak $s_W^2 = .01993$

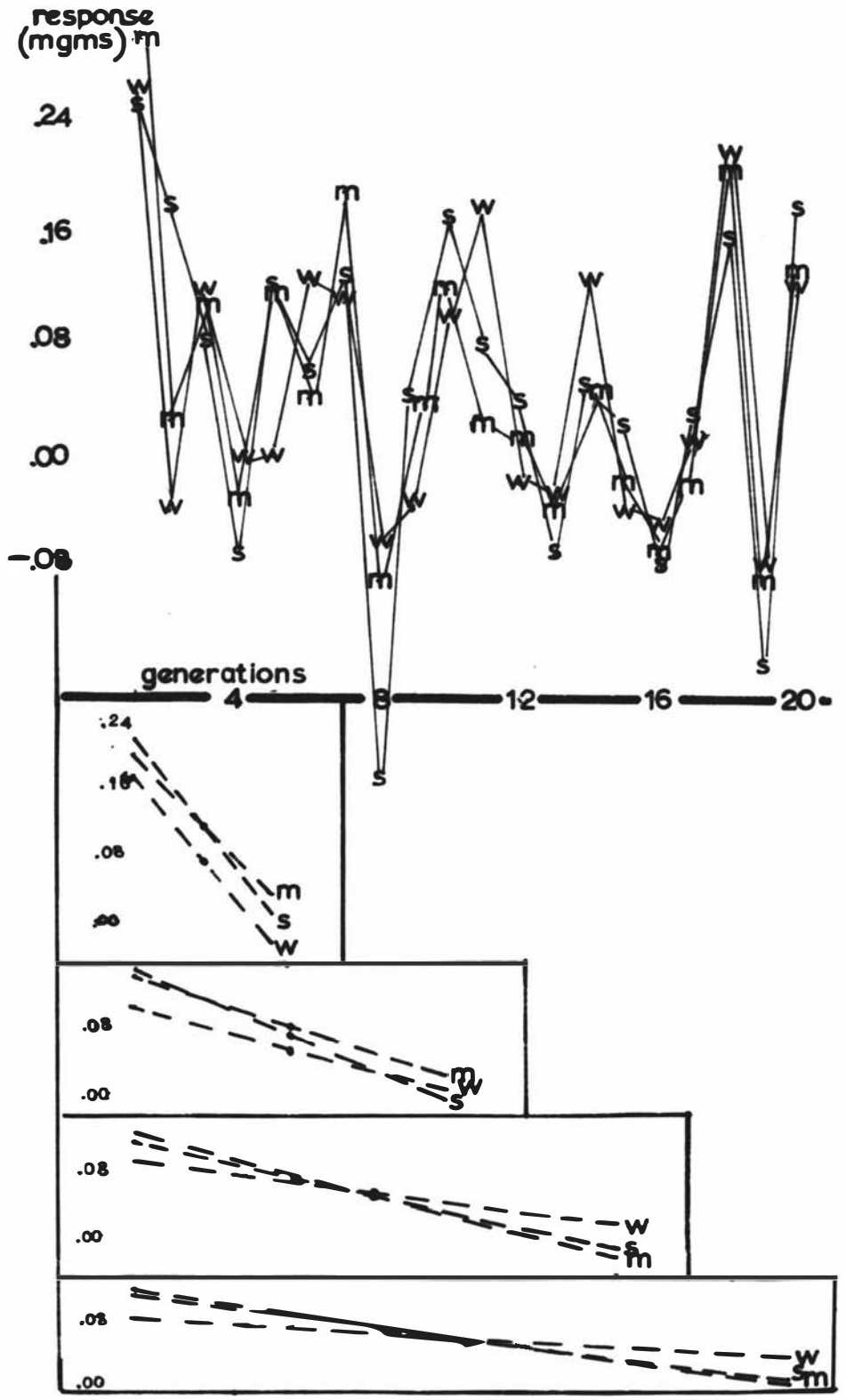
The F-ratios are $s_S^2/s_M^2 = 1.26$

$s_M^2/s_W^2 = 1.23$

The F-ratios are not significant on a two-tailed test at 6 x 19 = 114 and 114 degrees of freedom but it is interesting nevertheless to speculate on the pattern of results. The reason for such a pattern of differences to emerge between treatments cannot be established for certain. Lerner (1954) suggests that as inbreeding levels increase a population becomes more poorly buffered against the environment and therefore its response will fluctuate more greatly. This explanation would account for the largest values of standard deviation being those in Strong and the smallest values being those in Weak. It would also imply that such values should increase across the table as the length of period being considered increases and the inbreeding level consequently increases also within each treatment, but this expected effect is antagonised by the falling genetic variances in all treatments. The net result actually obtained is that the values are fairly constant with time in each treatment. The other explanation possible for these patterns is that the fluctuations represent sampling error as far as the choosing of parents is concerned. In the Strong treatment only two female and two male parents are selected each generation and this allows the possibility of considerable genetic drift, and therefore considerable fluctuation of response. In the Weak treatment eight females and eight males are chosen and so genetic drift would be very much less. This explanation accounts satisfactorily for the between-treatment patterns in Table 32 but it must still be regarded as tentative.

The standard deviations are very large compared with the means of Table 29 -

Fig 13 : FLUCTUATIONS AND TRENDS OF RESPONSE



they are much greater than their respective means on all occasions. This illustrates in a quantitative fashion the same conclusion that was made by the author from the graph patterns of Figure 13, namely that the fluctuations are very big indeed. There are a number of possible causes for high fluctuation and these include sampling error and weighing error. But one cause which should not affect the above estimates of standard error is fluctuation of the environment, because this factor is meant to be catered for by the Control population. However, suspicions of this type of Control for this type of experiment induced the author to re-calculate the standard deviations for each period, but this time without first adjusting the responses for variation of the Control population. The results of this calculation were quite disturbing. More often than not the values of standard deviation calculated in this way were smaller than their equivalents in the above table. In effect this means that response fluctuations (as measured by standard deviation) were not reduced by removing the Control variation, and this makes the author wonder about the value of such a Control. This point will be discussed later in the thesis but the result seems clear - if the fluctuations observed are partly due to changes in environment (and the previous experiment suggests that they were) then the Control has evidently not succeeded in its job of catering for them.

Regression coefficient: As explained earlier, regression lines were fitted to the five, ten, fifteen and twenty responses, each period in turn (Fig. 14). Then the six regression coefficients (one per replication) were averaged to give an estimate for each period. The validity of averaging the coefficients from all replications was checked for the twenty-generation period by an F-test of heterogeneity and the results are given below.

Fig 14 : REPLICATION DIFFERENCES

(a) PERIOD PARAMETERS

(generations 0 to 20)

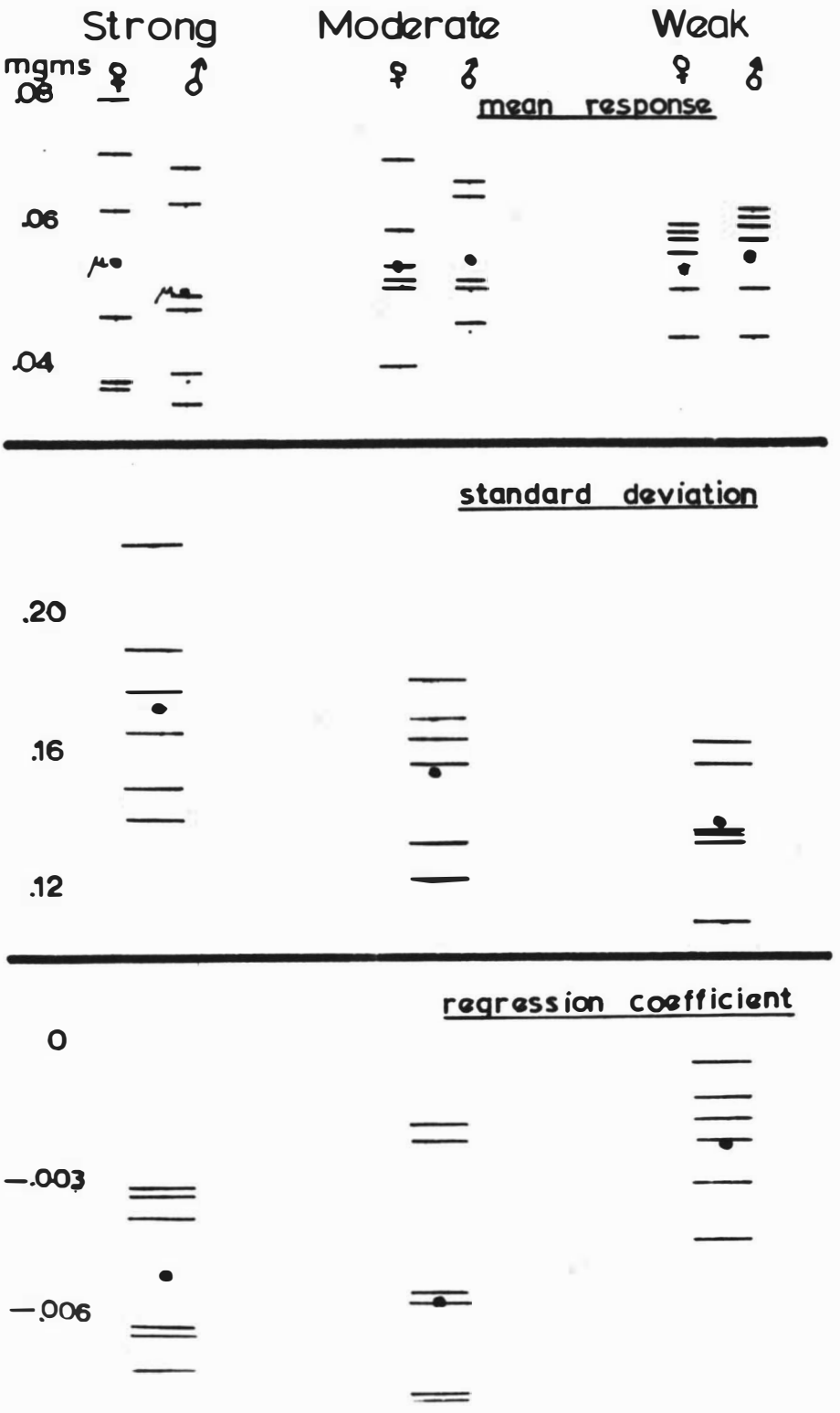


Table 33: Test for homogeneity of regression lines - Summary of F-tests

Source of variation	df	STRONG MS	STRONG F	MODERATE MS	MODERATE F	WEAK MS	WEAK F
Differences between regressions	5	.0019	0.07	.0023	0.11	.0030	0.16
Deviations from indiv. regressions	108	.0273		.0204		.0189	
Deviations from av. regressions within groups	113	.0262		.0196		.0182	

In all three treatments the F-ratios are far too small to indicate any trace of heterogeneity and so it is quite valid to pool the results of all six replications in order to obtain an average estimate of regression coefficient for the period. This done, the estimates were themselves compared between periods using the pooled mean squares of deviations (taken from the bottom line of the table above) and the standard t-test, for 113 degrees of freedom. On no occasion did this test approach a 5% level of significance, hence it must be concluded that the treatments do not show any great differences in the rate at which their selection response changes during the twenty generations of selection.

There are two more important points to be discussed about these regression coefficients. Firstly is the fact (Fig. 14) that in every replication of all three treatments over all four periods the regression coefficient has a negative sign. Regardless of whether each coefficient is itself significantly different from zero the consistency of their signs makes it fairly definite that the selection response per generation tends to decrease with time. As the period of selection becomes longer (left to right across Table 32) the value of this coefficient becomes closer to zero in each of the three treatments, which indicates

that although the selection response is at its greatest in the early generations it also tends to fall off most rapidly in the early generations. The second point of discussion arises from the extremely small F-values for heterogeneity of regression in Table 33. These F-values are so small that they indicate the six regression coefficients in any one treatment are much too similar in value than could be expected from the amount of deviation about each one. Instead of being heterogeneous the six values are in fact much too homogeneous. The most obvious interpretation of such a result is that a fluctuation of response among different generations is affecting all replications in similar fashion; for example, if there is a negative response in one generation for any given replication, the other replications will also tend to contain negative responses. This hypothesis was tested for the Strong treatment over the twenty-generation period and the relevant response data is given in Table 34.

Table 34: Correlated responses in Strong treatment

Gen	Rep 1	2	3	4	5	6	Total
1	.23	.32	.26	.34	.13	.20	1.48
2	.28	-.02	.22	.14	.30	.19	1.11
3	-.10	.10	.11	.20	.05	.09	.45
4	.08	-.09	-.03	-.18	-.07	-.11	-.40
5	.04	.38	.20	-.06	.00	.14	.70
6	.04	-.08	-.02	.17	.17	.06	.34
7	-.24	.12	.24	.14	.19	.32	.77
8	-.01	-.15	-.31	-.14	-.41	-.34	-1.36
9	-.06	-.11	.04	.07	.18	.11	.23
10	.11	.38	.04	.07	.18	.24	1.02
11	.04	.19	.15	.07	.03	.03	.51
12	-.05	.18	.18	.04	-.02	-.10	.23
13	-.02	-.11	-.14	-.07	-.18	.08	-.44
14	.07	.00	.03	.10	.12	-.02	.30
15	.11	.10	.10	-.01	-.08	-.07	.15
16	-.03	-.17	-.19	.11	-.04	-.17	-.49
17	.00	.11	-.04	.12	-.01	.02	.20
18	.17	.02	.31	-.08	.17	.38	.97
19	-.13	-.12	-.11	.02	-.13	-.44	-.91
20	.10	.34	.08	.18	.09	.30	1.09
Total	.63	1.39	1.12	1.23	.67	.91	5.95

An analysis of variance was carried out on this data, with the following result:

Source of variance	df	MS	F	P
Replications	5			
Generations	19	.087832	5.99	.001
Residual	95	.014651		
Total	119			

The extremely significant F-value illustrates statistically what can be seen visually in the table, namely that there is a marked tendency for the six replications to vary in unison. The only factor powerful enough to cause this effect was, in the author's opinion, fluctuation of the environment. It is quite reasonable and indeed to be expected that replications which are fairly similar genetically will respond in fairly similar fashion to changes in the environment. What is not reasonable or expected however, is that this effect should persist in the data even after adjustment has been made for variation of the Control population; for the main purpose of such a Control was to cancel out this effect. The author regards this result as additional evidence that his Control was not satisfactory for this type of experiment and the matter will be discussed further later.

(b) DIFFERENCES BETWEEN REPLICATIONS:

When variability between replications is measured by a heterogeneity analysis and expressed in terms of a statistical probability it does not always give the reader a very concrete vision of variability. The alternative method used in the earlier experiment would appear to be appropriate here also; that is, for each of the three response parameters in turn (mean, standard deviation, regression coefficient) calculate the standard deviation of the six replication values about their mean value. These standard deviations can then be divided by their respective mean values to give a coefficient of variation for each parameter. All four periods of response have been analysed in this way for Table 35.

Table 35: Replication differences a) period parameters

Generations	0 to 5		0 to 10		0 to 15		0 to 20	
	S.D.	C.V.	S.D.	C.V.	S.D.	C.V.	S.D.	C.V.
	a) mean response							
Strong	.0292	27%	.0182	26%	.0173	31%	.0143	28%
Moderate	.0171	16%	.0196	24%	.0101	18%	.0087	16%
Weak	.0162	24%	.0094	17%	.0048	9%	.0066	12%
	b) standard deviation							
Strong	.0486	28%	.0154	8%	.0140	8%	.0280	16%
Moderate	.0293	19%	.0183	12%	.0157	10%	.0219	14%
Weak	.0412	26%	.0313	23%	.0274	20%	.0182	13%
	c) regression coefficient							
Strong	.0405	79%	.0094	57%	.0028	30%	.0018	36%
Moderate	.0322	77%	.0112	89%	.0022	20%	.0019	34%
Weak	.0317	67%	.0059	54%	.0041	73%	.0012	52%

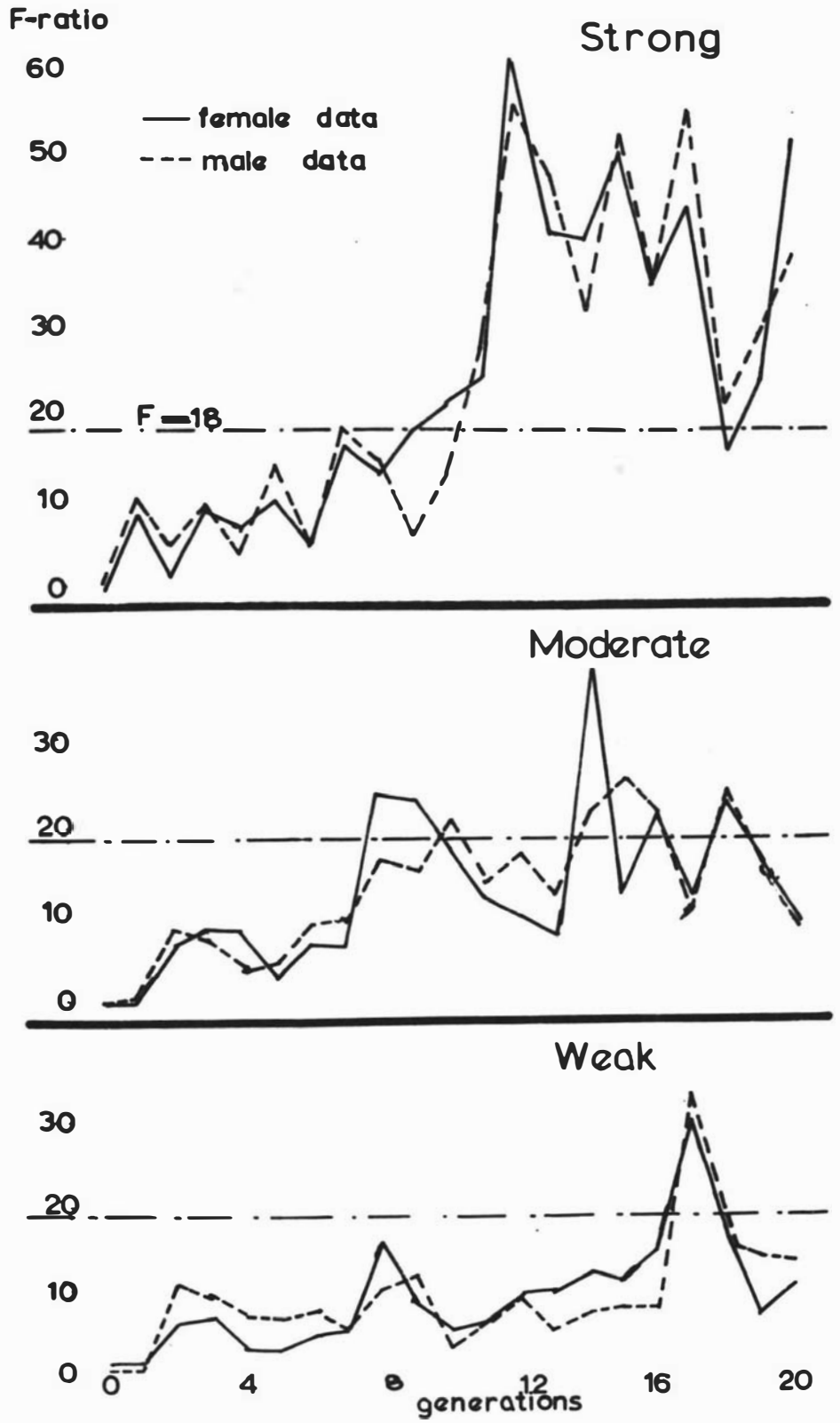
It is unwise to look too deeply for patterns between treatments or between periods in this table for it is difficult to measure the significance of such patterns. There is in fact apparently only one such pattern in the whole table, namely the tendency for variability to decrease from Strong to Moderate to Weak in the parameter of mean response. To the extent that this pattern is significant it may be caused by differential amounts of genetic drift, which would theoretically be greatest in the highly inbreeding Strong treatment and least in the Weak one. More refined evidence of this pattern will be given by a comparison of variance components in the following section.

The average coefficient of variation of replications about their mean is about 21% for the mean response, 16% for standard deviation and about 56% for regression coefficient. As was the case in the first experiment regression coefficient is most variable when measured by this method; the fact that in this second experiment its coefficient of variation is much lower than in the first one is usually because the denominator of this coefficient is not so close to zero.

Variability between replications can be measured each generation instead of over a period. This time it is the mean weights of replications that are compared rather than their responses. The variability between them is now expressed as an F-ratio, between-replication variance to within-replication variance, and the value of this ratio is graphed in Figure 15. If this ratio rises with time it indicates that the differences between replications are becoming more significant compared with the overall population variability. The ratio would appear to be free of bias from the effects of scale and natural selection.

Each of the graphs in Figure 15 shows at least one large fluctuation but nevertheless they all contain quite clear patterns of response. There is an

Fig 15 : REPLICATION DIFFERENCES
(b) VARIANCE RATIO



increase of F-ratio quite regularly through the twenty generations of selection and this ratio passes the 5% (1%) significance value of 2.3 (3.2) within the first one or two generations. As shown by the horizontal guideline at F = 18 the increase is most rapid in Strong and least rapid in Weak. Once again this difference between treatments is explainable in terms of genetic drift. Those treatments in which fewest parents are chosen each generation will be most likely to produce chance genetic differences between replications and these differences will cumulate as generations pass by. On this argument it is not unexpected that Strong should show the greatest amount of variability, but the absolute largeness of these F-ratios in all three treatments is perhaps surprising. The sudden rise in F-ratio for Strong at about Generation 12 is possibly a reflection of the very great divergence of Replication 1 (Fig. 10) rather than a reflection of increased divergence among all replications and if this is the case it may be that a genetic accident such as mutation has affected Replication 1.

Genetic drift is not the only cause of variability between replications; another cause is the effect of artifical selection operating on these differences created by drift. The relative effects of these two forces (drift and selection) is measurable but it must be stressed that the method of measurement is only approximate. This method consists of calculating the value of the between-replication variance component which would be expected if only genetic drift was present and no selection was being applied. Subtraction of this theoretical value from the observed value of between-replication component leaves an amount which can be attributed to the differential effects of selection.

The amount of variability expected from drift is given by

$$\sigma^2 \text{ between reps} = 2F\,\sigma^2 \text{ add}$$

where F = inbreeding level

σ^2 add = additive variance in base population.

The values of additive variance for female and male data were obtained from the between-progenies component of a fullsib analysis of variance. This variance analysis was carried out as part of a test of prediction theory and its formulation is described later in the experiment; meanwhile it produced the following estimates of additive variance.

female data: $\sigma^2_{add} = .01260$

male data: $\sigma^2_{add} = .01820$

It is necessary also to calculate the level and rate of inbreeding at all generations. The circular types of mating system in the Strong, Moderate and Weak replications are virtually equivalent to double first cousin, quadruple second cousin and octuple third cousin mating respectively. The inbreeding level therefore rises each generation according to the formulae of Wright (1921a)

$$F = \tfrac{1}{8}\,(4F' + 2F'' + F''' + 1) \text{ for Strong}$$

$$F = \tfrac{1}{16}\,(8F' + 4F'' + 2F''' + F'''' + 1) \text{ for Moderate}$$

$$F = \tfrac{1}{32}\,(16F' + 8F'' + 4F''' + 2F'''' + F''''' + 1) \text{ for Weak.}$$

In each case the inbreeding rates quickly settle down to an almost constant value and so the mean inbreeding rate over all generations has been adopted, namely $\Delta F = .090$ for Strong, .038 for Moderate, and .018 for Weak - which correspond respectively to the inbreeding rates achieved by random mating of 5.6, 13.2 and 27.8 parents. These latter values have now been modified slightly to cater for the effect of selection within each fullsib family, by the formula of Robertson (1961).

$$N/N_e = (1 - \tfrac{1}{2n}) + 4C^2\,(1 - \tfrac{1}{n})$$

where n = number of animals per sex weighed in each family

C^2 = variance of relative selective advantage between families.

In the present experiment one female and one male were selected from each fullsib family and equal numbers of animals were weighed in each family (sixteen

of each sex weighed in Strong, eight of each sex in Moderate, four of each sex in Weak). Therefore C = 0 and effective population size is given by

$$N_e = 5.6 \times 32/31 = 5.8 \text{ for Strong}$$

$$N_e = 13.2 \times 16/15 = 14.1 \text{ for Moderate}$$

$$N_e = 27.8 \times 8/7 = 31.8 \text{ for Weak.}$$

The corrected rates of inbreeding ($1/2N_e$) are therefore

$$\Delta F = .086 \text{ for Strong}$$

$$\Delta F = .035 \text{ for Moderate}$$

$$\Delta F = .016 \text{ for Weak}$$

The inbreeding level at any generation t is then calculated as

$$F_t = \Delta F + (1 - \Delta F)F_{t-1}$$

and the final step in estimating the value of variance due to genetic drift is by use of the formula

$$\sigma^2 \text{ between reps} = 2F\ \sigma^2 \text{ add}$$

Table 36 contains for each treatment both of these results - the inbreeding level at each generation and estimated from it, the variance due to drift for each sex.

The variance data of this table are presented again as a graph in Figure 16 to enable comparison with the between-replication variance data actually recorded in the experiment. It is judged from this figure that in all three treatments the variance due to drift is again less than half of the total variance between replications, just as was the case in the first experiment. The remainder is apparently due to differential effects of selection acting on these differences between replications set up by genetic drift.

Fig 16: REPLICATION DIFFERENCES

(c) BETWEEN-REPLICATION VARIANCE COMPONENT

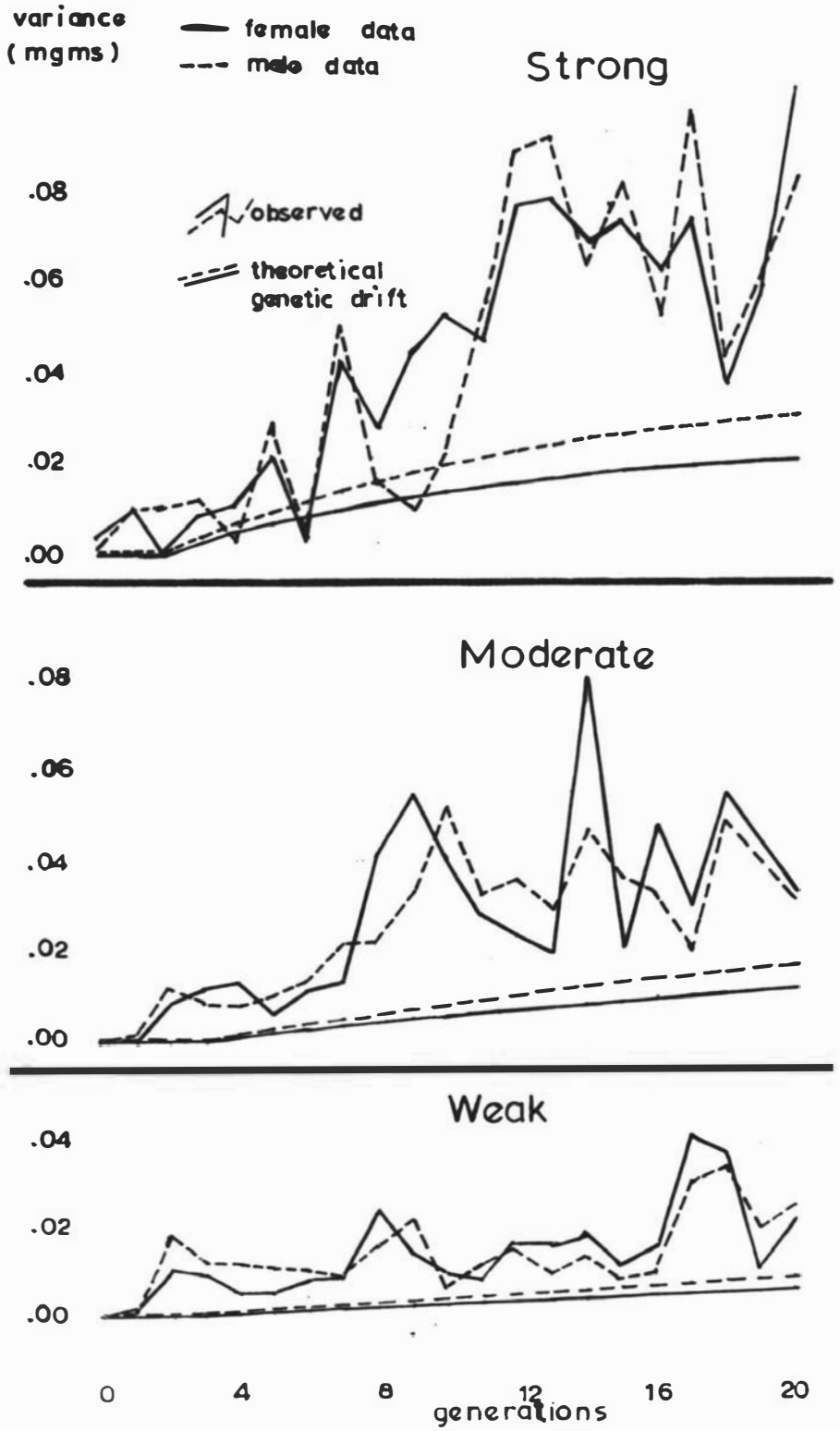


Table 36: Replication differences

b) between-replication variance component

	STRONG			MODERATE			WEAK		
	F	variance between reps		F	variance between reps		F	variance between reps	
Gen		female data	male data		female data	male data		female data	male data
0	.000	.0000	.0000	.000	.0000	.0000	.000	.0000	.0000
1	.000	.0000	.0000	.000	.0000	.0000	.000	.0000	.0000
2	.000	.0000	.0000	.000	.0000	.0000	.000	.0000	.0000
3	.086	.0022	.0031	.000	.0000	.0000	.000	.0000	.0000
4	.165	.0042	.0060	.035	.0009	.0013	.016	.0004	.0006
5	.237	.0060	.0086	.069	.0017	.0025	.032	.0008	.0012
6	.303	.0076	.0110	.102	.0026	.0037	.047	.0012	.0017
7	.363	.0091	.0132	.133	.0034	.0048	.062	.0016	.0023
8	.418	.0105	.0152	.163	.0041	.0059	.077	.0019	.0028
9	.468	.0118	.0170	.192	.0048	.0070	.092	.0023	.0033
10	.514	.0130	.0187	.220	.0055	.0080	.107	.0027	.0039
11	.556	.0140	.0202	.247	.0062	.0090	.121	.0030	.0044
12	.594	.0150	.0216	.273	.0069	.0099	.135	.0034	.0049
13	.629	.0159	.0229	.298	.0075	.0108	.149	.0038	.0054
14	.661	.0167	.0241	.323	.0081	.0117	.163	.0041	.0059
15	.690	.0174	.0251	.347	.0087	.0126	.176	.0044	.0064
16	.717	.0181	.0261	.370	.0093	.0135	.189	.0048	.0069
17	.741	.0187	.0270	.392	.0099	.0143	.202	.0051	.0074
18	.763	.0192	.0278	.413	.0104	.0150	.215	.0054	.0078
19	.783	.0197	.0285	.434	.0109	.0158	.228	.0057	.0083
20	.802	.0202	.0292	.454	.0114	.0165	.240	.0060	.0087

3.32 Predicting selection response

Prediction from base population: When this second experiment was originally planned it was intended that the test of prediction theory would take exactly the same form as that test applied in the first experiment. That is, several methods of predicting a response by variance analysis of the base population would be examined, followed by a study of the first-generation response as an alternative predictor. This original plan had to be modified because of conditions which developed during the experiment. As both experiments neared their end certain populations in each of them became very slow at producing their quota of progeny and when one of these populations was submitted to a pilot diallel analysis there were very many failures caused by death or low fertility of parents. Because of this high failure rate many more 2 sire - 2 dam mating blocks would have to be set up than was originally planned if the resulting estimates of heritability were to have low standard errors, and these mating blocks would have to be kept far longer than originally expected. This situation called for far more incubator space than was anticipated, or available. It left the author with two alternatives - either he could attempt to carry out equal diallel analyses on both experiments but with reduced numbers and degrees of freedom in each, or he could ensure that one experiment be given a complete and precise diallel analysis and the other experiment be analysed in a less sophisticated manner. The author chose the latter alternative on the grounds that one precise test of prediction is better than two imprecise ones, and he chose the first experiment to contain the diallel analysis in the belief that its selected stocks would be fitter and therefore more likely to complete such a diallel programme. This decision meant that the prediction of heritability in Experiment Two must suffer accordingly a reduction in the size of analysis, and it was decided to replace the intended diallel system by a fullsib analysis. Accordingly female and male parents were chosen randomly

from the base population and mated in pairs. A random four female and four male pupae of the progeny of each pair were sampled and weighed, these weights providing the raw data of the fullsib analysis. The same fullsib mating experiment was also carried out on the selected Strong, Moderate and Weak populations at halfway (Generation 11) and at the end (Generation 20) of the experiment, as it was felt that these extra analyses might show some patterns of changing genetic variance and so partly compensate for the loss of sophistication at the base population. Finally it is obvious that this fullsib type of analysis does not permit either halfsib estimates of heritability or Abplanalp's linear estimates of heritability. The remaining two types, fullsib estimates and parent-offspring regression estimates are given in Table 37, along with the measurements of realised heritability for the generation concerned, and for the mean of five generations either before or after it.

The following comments can be made about Table 37.

(1) Realised heritability: Previous observation of the very big fluctuations in realised heritability should caution against placing too much emphasis on the results from just one generation. The single-generation results given in the table do not show very much relation to the five-generation results given below them and it is considered by the author that the latter are a more reliable indication of genetic variance remaining in the population. This attitude is based partly on the fact that fluctuations of the environment have less overall effect during five generations than they have during one generation, and until the Control population shows more ability to cater for these environmental fluctuations the realised heritability is better considered in terms of average value over a period. It is therefore these five-generation measures of realised heritability which are used as the basis of deciding if prediction has been successful.

(2) Fullsib analysis: The fullsib estimates of heritability are far too

Table 37: Estimation of Heritability

a) Base population

	d.f.	Mean Square	
Full-sib analysis		female progeny	male progeny
Between progenies	69	.07696	.07957
Within progenies	210	.02655	.00676
h^2 =		.64 ± .13	1.46 ± .06

Regression analysis (90 progenies, each of 4 females and 4 males)

regr. on dam	h^2 =	.33 ± .06
regr. on sire	h^2 =	.39 ± .05

Realised heritability : Generation 1 (average of 6 reps, each of 32 pupae)

	♀	♂		♀	♂		♀	♂
Strong selection h^2 =	.59	.51	Moderate	.89	.82	Weak	.93	.90
: Generations 1 to 5	.33	.02		.39	.30		.25	.23

b) Generation 11

		Strong			Moderate			Weak	
Full-sib analysis		♀	♂		♀	♂		♀	♂
Between progenies	32	.24562	.18187	33	.11364	.14939	34	.10824	.12412
Within progenies	99	.03596	.05495	102	.05902	.05589	105	.04543	.04114
h^2 =		1.19 ± .16	.73 ± .19		.38 ± .18	.59 ± .19		.51 ± .18	.67 ± .19
Regression analysis		33 progenies			34 progenies			35 progenies	
regr. on dam h^2 =		.94 ± .07			.47 ± .08			.54 ± .07	
regr. on sire h^2 =		.88 ± .09			.30 ± .08			.13 ± .09	
Realised heritability		6 reps			6 reps			6 reps	
: Generation 11 h^2 =		.26	.12		.01	.16		.84	.73
: Generations 10 to 14		.11	.09		.11	.07		.38	.31

c) Generation 20

		Strong			Moderate			Weak	
Full-sib analysis									
Between progenies	30	.38600	.28000	30	.19800	.37933	34	.20971	.20235
Within progenies	93	.09355	.06032	93	.06688	.07688	105	.07867	.07743
h^2 =		.88 ± .19	.95 ± .19		.66 ± .20	.99 ± .19		.59 ± .19	.57 ± .19
Regression analysis		31 progenies			31 progenies			35 progenies	
regr. on dam h^2 =		.69 ± .13			.71 ± .12			.48 ± .11	
regr. on sire h^2 =		.29 ± .14			.14 ± .12			.18 ± .10	
Realised heritability		6 reps			6 reps			6 reps	
: Generation 20 h^2 =		.59	.40		.30	.42		.55	.68
: Generations 16 to 20		.06	.06		.02	.11		.23	.33

high in the base population and they show little tendency to decline during selection. This pattern is strikingly similar to that obtained from fullsib analysis in the first experiment and it is likely that the same explanations would apply, namely that the estimate is biased upwards by large amounts of non-additive genetic variance and by other effects (e.g. maternal) common to fullsib litters.

(3) Parent-offspring regression analysis: Estimates of heritability by regression analysis in the base population are quite similar to those for realised heritability. At the halfway and endpoint analyses this heritability prediction is still similar to realised heritability when the prediction is based on offspring-sire regression (except for the one discrepant - Strong at Generation 11) but it is far too inflated when based on offspring-dam regression. Once again these results are very much like those from the first experiment and again this difference between offspring-sire and offspring-dam estimates is attributed to a strong maternal effect.

(4) If the explanations given above are accepted it can be claimed that offspring-sire regression analysis, and only this one of the methods tested, will provide a reasonably accurate prediction of the heritability for populations like that used here. But neither this system nor any of the others adequately predicts that Strong heritability will fall faster than Moderate or Weak, as is expected in theory, and realised in the results of the experiment. From this point of view there is little merit in predicting the heritability of a population by fullsib or regression analysis.

Prediction from first-generation response: The value of selection response at the first generation for predicting subsequent response is examined in the same way as for the previous experiment, that is by calculating the correlation coefficient between the two parameters. Within each treatment the value of first-generation response is paired with the value of mean response for that period, each replication in turn. Then the correlation coefficient is calculated among the six pairs. This same process is carried out with the standard deviations and regression coefficients of the period, and then with all three parameters of the longer periods as well. The resulting values of correlation coefficient are given in Table 38.

Table 38: Correlation between first-generation response and other parameters

Generations	0 to 5	0 to 10	0 to 15	0 to 20
	r(first:mean)			
Strong	.16	.04	.59	.81
Moderate	.19	.31	.22	.33
Weak	.15	.57	.16	[illegible]
	r(first:standard deviation)			
Strong	.80	-.08	-.22	-.26
Moderate	.71	-.12	-.27	-.63
Weak	.89	.85	.90	.88
	r(first:regression coefficient)			
Strong	-.66	-.83	.17	.57
Moderate	-.56	-.60	-.71	-.15
Weak	-.91	-.45	-.46	-.41

The correlation between first-generation response and mean response is always positive but never significant. The former response therefore cannot be regarded in this experiment as a very precise predictor of the latter response. The statistical significance of these correlations might increase in more precise and well-replicated experiments but this does not imply that the correlation values themselves would be ever high enough to use for prediction.

The correlation between first-generation response and standard deviation is often large in a positive direction and often small in a negative direction. The positive bias of these results is again suggestive of a real correlation between the two parameters but it is too inconsistent in the present experiment to be useful for predicting the fluctuation of response.

The correlations between first-generation response and regression coefficient have a negative value in ten of the twelve cases. The interpretation of this trend is that as the value of the former increases the value of the latter becomes more negative; in non-statistical language it means that the greater is a response in the first generation the more rapidly it will fall off in later generations. This is entirely consistent with additive genetic theory, for both parameters depend on the amount of genetic variance present in a population. Once again however, use of the one to predict the other would seem to require high replication and big populations.

The reason for discrepancy between first-generation response and mean response is worthwhile investigating, just as in the previous experiment. The discrepancy between them has been calculated for each replication in turn and from this the average value (disregarding the sign) and the standard deviation of these discrepancies. The results are given in Table 39, firstly for the case when the replications are kept entirely separate and also no adjustment

is made for variation of the Control, secondly for when this adjustment for Control is made beforehand (requiring also adjustment of the mean), thirdly for when no adjustment is made but the responses of all replications are pooled beforehand, fourthly for when both adjustment and pooling take place before the discrepancies are calculated.

Table 39: Discrepancy between first-generation response and mean response

Generations	0 to 5		0 to 10		0 to 15		0 to 20	
	average discrep.	S.D.	average discrep.	S.D.	average discrep.	S.D.	average discrep.	S.D.
Strong								
Control and reps both unpooled	.11	.072	.17	.077	.19	.072	.19	.071
Control unpooled, reps pooled	.11		.17		.19		.19	
Control pooled, reps unpooled	.14	.075	.18	.081	.19	.071	.19	.070
Control and reps both pooled	.14		.18		.19		.19	
Moderate								
Control and reps both unpooled	.17	.067	.22	.075	.25	.071	.25	.070
Control unpooled, reps pooled	.17		.22		.25		.25	
Control pooled, reps unpooled	.20	.063	.23	.074	.25	.067	.26	.070
Control and reps both pooled	.20		.23		.25		.26	
Weak								
Control and reps both unpooled	.16	.046	.19	.042	.20	.043	.19	.046
Control unpooled, reps pooled	.16		.19		.20		.19	
Control pooled, reps unpooled	.19	.046	.20	.040	.20	.045	.20	.046
Control and reps both pooled	.19		.20		.20		.20	

The average discrepancies between first-generation and mean responses are all except once very much larger than the respective mean responses themselves recorded back in Table 29. Furthermore, the discrepancies tend to increase in value with

the length of period being considered, though this pattern merely reflects that the mean response itself is decreasing as time goes by. Thirdly there is no improvement made in the discrepancy by pooling all six replications before it is calculated, and the fact that the pooled and unpooled values are always identical is because the first-generation response is always greater than the mean response and therefore the discrepancy is always positive in sign in all replications.

The standard deviations are also very large when compared with the mean responses of Table 29 and this means that there is a large range in the size of discrepancy. In two senses therefore the first-generation response is a poor predictor of the period mean response: firstly the replications vary considerably in their discrepancy between the two; secondly the discrepancy is still very large even when the replications are pooled.

There is one other important result in Table 39, namely that the average discrepancy actually increases on most occasions when adjustment is made for the Control response. This can only be regarded as further evidence that the type of Control used was not a very suitable one for this type of experiment.

Finally, the standard deviation of discrepancy decreases systematically from Strong to Moderate to Weak in all four periods, even though the average discrepancies themselves show no such pattern. This is yet another manifestation of the lesser variability between Weak replications than between Moderate or Strong ones and it probably just means that genetic drift is least in the Weak treatment, as is expected from genetic theory.

It does not appear likely that the limit of selection response had been reached in this experiment when it closed at Generation 20. Therefore no attempt was made to apply the equation of Robertson (1960) for predicting the ratio of total response to first-generation response.

3.4 DISCUSSION

Occasionally through this experiment the author has commented that the Control was not doing its job. The main reasons for this comment were:

(a) it did not reduce the fluctuations, as measured by standard deviation between generations;

(b) it caused the regression coefficients to be more alike than they should be;

(c) it did not reduce the discrepancy between first-generation response and mean response.

From a theoretical viewpoint one can of course not expect a single Control to cater adequately for all three treatments because immediately these three treatments begin to diverge in response they are being affected differently by such factors as scale, natural selection and perhaps genotype-environment interaction. But if such a Control is not expected to stay close to any one treatment in genetic content it is alternatively surely expected to stay itself constant in genetic content, otherwise it will not even record a constant reaction to changes in the environment. These two requirements were laid down by Bray et al., (1962) in their study of Control populations for Tribolium and they specified that for such requirements to be met the Control population should contain at least fifty pairs of parents each generation and that mass-mating should be not used. These specifications were followed in the present experiment and accordingly the Control should be valid in the sense of reflecting changes in the environment. The author is forced to conclude that such changes in the environment affected the three treatments in an entirely different way, so different in fact that it would be better to not adjust for this Control at all. The author did indeed re-calculate the results on this new basis of ignoring the Control but they are not given here because they do not usually affect the comparison between treatments at all, nor do they appreciably alter the accuracy of predicting response.

Regardless of the value of Control then, the comparison of Strong, Moderate

and Weak selection intensities contains results that should be of value to commercial breeders. It shows that the responses to selection at the three intensities did indeed rank in the correct theoretical order; Strong was highest at the start, Moderate was highest at the middle, and Weak was highest at the end. However, the differences between them were not very significant and until such is the case it implies that the choice of selection intensity for a breeding project should be dictated by criteria other than just expected response. For example Strong declined greatly in fitness toward the end of the experiment; whether or not it was due to the high inbreeding rate this weakness would be a serious set-back in commercial projects. From the limited results of the present experiment the most suitable selection intensity would be Moderate, for over all four periods it was never more than slightly below the best in response and it did not appear to suffer from a loss of fitness. This conclusion is at least an advance on Lerner's opinion (1958, page 123) that "the deficiency in current selection theory makes it necessary to follow rule of thumb in deciding on the selection intensity to be used in any given situation".

GENERAL DISCUSSION

During this thesis the author recorded subsidiary data on characters such as sex ratio in each replication and the tendency of replications to hold their rank. No use of this data has been made at present because it would break continuity of the main theme. However, a close watch was kept on the generation interval because this parameter usually reflects any fitness changes in a line. The weaker a line becomes the longer it will take to produce its quota of 32 female and 32 male progeny. In Experiment One there was a slight loss of fitness in the Heavy population but a much greater one in the Light population. In the closing generations up to seven weeks was needed instead of the original four weeks, for a generation. The Strong population of Experiment Two suffered similarly but probably because of inbreeding effects rather than animal size. This failing required that the Moderate and Weak lines had to be held back also, for all three treatments and the Control must be kept in step. Also, some of the weakest replications had to be sampled more than once before they produced the quota of progeny.

However, all the extra work that this falling fitness brought about, including the carrying of two extra replications, had one big advantage. It meant that the author was able to avoid having big gaps in the data, which would have weakened the analysis of results. This seems to be a serious weakness in the work of Sheldon (1963) and DeFries and Touchberry (1961) and some other authors, but in the present experiment the full number of replications and of animals in each one was maintained to the very end. Perhaps its biggest weakness is that there were not more replications and more animals in each one from the start. If this had been the case the fluctuations between generations would probably have been far smaller than they were. Evidence for this is given in the table below. The data listed are the values of standard deviation between generations for Experiment

One. Firstly it is calculated separately for each replication and the six values are averaged. Secondly the responses of all six replications are pooled for each generation in turn, before calculating the standard deviation. The effect of this second procedure should be analogous to increasing the population size, though not necessarily six-fold. (On the one hand the higher inbreeding rates in the separate replications might reduce their ability to buffer against changes in environment but on the other hand the effects of genetic drift may partially cancel one another when the replications are averaged in each generation. The first factor operates to increase standard deviation, the second one operates to reduce it).

Table 40: Effect of population size on fluctuations

Generations	1 to 6	7 to 12	13 to 18
		Heavy population	
reps separate	.0635	.0475	.0550
reps pooled	.0227	.0199	.0147
		Light population	
reps separate	.0732	.0565	.0409
reps pooled	.0337	.0202	.0180

On all occasions the standard deviation is greatly reduced when the replications are pooled in each generation beforehand. In the same experiment it was recorded earlier that the discrepancy between first-generation and mean responses was mostly reduced to a reasonably low value if all replications (and their Controls) were pooled before the discrepancy was calculated. Another benefit that

would have been obtained by increasing population size involves the number of parents selected - the author would have been able to select more parents in each generation while still maintaining the desired selection intensity. As it was, there were only four parents per replication in the Strong treatment and so genetic drift effects may have been quite large. They may in fact be one reason that additive genetic theory was sometimes not supported.

For the most part the results of this thesis followed additive genetic theory, but never closely enough to use for prediction. The reasons for this were firstly the fluctuations discussed earlier and secondly the presence of maternalism and asymmetrical responses. The overall impression gained is that by merely increasing population size and ensuring a reliable Control, fairly accurate predictions using the first-generation response could be made. Even in the present experiment there were often clear associations between this first-generation response and the mean response of the period following it. Future experiments could show that the association is reliable enough to predict from.

The first suggestion for future work is therefore a similar test of genetic theory using a range of population sizes. Once this sort of basic data is on hand students could then study more sophisticated topics such as the effect of changing selection intensity part-way through an experiment. Perhaps there is even greater knowledge to be gained from Fraser's approach. He suggests (pers. comm.) the technique of restricting polymorphism to specified small chromosomal segments, in order to follow the effect of selection acting on inter- or intra-segment variability. At present this would be possible only in organisms well charted genetically, such as Drosophila. In lesser-known species such as Tribolium work would be limited to imposing bottlenecks of various sorts, in order to create a range of disequilibria in the base population.

It seems likely that computers will be used much more frequently in future to interpret selection responses. Indeed it is possible that some past experiments

will be resurrected for computer analysis, as was done by Fraser for Mather's work. The present thesis should have much more value when it is simulated on a computer eventually and the results can be analysed further. But it should be remembered that a computer can never be as good a test of genetic theory as the experiment itself, and in turn the value of this theory for commercial breeders can be reliably determined only from the commercial populations themselves. This implies that such breeders are under obligation to sacrifice some profit for the sake of testing genetic theory. For example, they should carry more refined Controls, should operate more frequent genetic analyses, and should continue selection for longer than is profitable. Only in this way can the theoreticians be guided as to where their genetic models are faulty or poorly adapted to help the commercial breeders. It seems to the author also that statisticians could contribute more thought to the needs of an experimenter. Too often in the present thesis was there no reliable significance test for the results, for example variability between replications over a period. Even the diallel analyses were strictly speaking not meant to be applied to populations that had been selected for so long. The reason they were applied was simply that no better tests were available.

The overall view is perhaps best expressed by Robinson (1963). He is summing up the results of a recent conference set up specifically to evaluate statistical genetic theory in its application to plant breeding. He acknowledges that this theory is still deficient, especially in regard to epistasis, linkage and genotype-environment interactions. Some of the statistical procedures have limited population inferences, for example diallel analyses. This means that the theory should be tested on every possible occasion by experiments that are well constructed and well analysed. In Robinson's words (page 433) - "As has been repeatedly demonstrated in this conference, it is the joint efforts of experimentation and theory proceeding on an integrated basis that have contributed

to significant progress in our knowledge during the past 15 years". It is hoped that the present thesis may have some value in relating the two, in bringing about this wedding of unproven theory and uneducated practice. The author regards his overall thesis results as supporting genetic theory mostly only in a qualitative sense. Certainly the agreement was not precise enough for a breeder to confidently predict a selection response on the basis of such theory. The author has two comments to make about this result, firstly that the prediction failures could usually be explained or speculated away in terms of scale, non-additive gene action etc, which implies that they might be catered for in advance with extremely sophisticated experiments. Secondly there were adequate signs that predicting from a first-generation response, detecting significant differences between Strong, Moderate and Weak selection etc, might be successfully achieved by greatly enlarging the size of population, number of replications, and by ensuring a good Control. That is, it seems likely that selection theory would be proved to be correct and applicable, but only in an "ideal" experiment of almost infinite size and detail. From a commercial breeder's viewpoint this finesse is just not possible and the results of the thesis suggest that he would be better to ignore such standard techniques as predicting from a diallel analysis. It seems to the present author that since this thesis experiment was far more sophisticated than a commercial breeder's projects would be, and that yet this thesis gave mostly only fair support to analytical and predictive theory, then such selection theory is drastically out of touch with the needs of a commercial breeder. Whether it is incumbent upon theoreticians to take a more down-to-earth attitude to their work, or incumbent upon breeders to become less concerned about profits and more concerned about understanding what they do, is not the author's privilege to say.

On a more positive note, perhaps the strongest point of this thesis is its high and consistent replication. Most of the statistical significance tests applied were based on the variability between replications and it seems clear

that many more treatment differences would have become significant had more replications been carried. As was the case for DeFries and Touchberry (1961) and Marion (1958) the replications diverged considerably right from the start. Observation of the graphed responses in Figure 1 and Figure 10 showed that after such divergence however the replications tended to hold their ranking order. This explains why there is a very high between-replication variance component and yet little heterogeneity between replications in the sizes of their fluctuations and trends in each period.

Finally the usual warning must be made about extrapolating the results of such a thesis. Robertson (1958) discusses this problem in detail and warns that it is very dangerous to extrapolate from one population to another, even on the statistical level, without first understanding the genetic situation in both populations. Experience of the literature indicates that it is more reliable to extrapolate from one species to another for the same type of selection character than to extrapolate from one character to another within a single species. This is a particularly important principle if the selection characters concerned affect the fitness of an animal. Characters such as bristle number in Drosophila do not greatly affect the animal's ability to survive and breed in laboratories and it has been observed that such characters generally contain much additive genetic variance and little inbreeding depression and therefore a high heritability of response to selection. But characters such as egg-laying ability are very closely connected with Drosophila's welfare and in such cases a shift by artificial selection in the population mean value is usually accompanied by natural selection against further progress. There is usually little additive variance present but considerable non-additive variance, and this is understandable since natural selection would much earlier have used up all of the additive variance available in shifting the population to its original equilibrium position. The

present selection character in Tribolium (body weight) appears to have little drastic effect on fitness until late in the experiment, by which time much response has already been obtained, and the author sees no obvious reason why similar results should not be obtained for body weight selection on comparable populations of sheep or plants. However, the word "comparable" is extremely important in this claim because the present thesis has shown that selection response is affected by many other factors besides the type of selection character - factors such as previous selection history, changes in the environment, and size of population. Such factors as these are the basis for Hanson's (1963) plea that whenever an author reports the heritability of a selection response he should also report the conditions in which the heritability value was calculated - such as the type and size of population, number of replications, type of environment, mating systems etc. The benefit of the present thesis to any other breeder thus depends on the degree of similarity between these conditions and his own. If his conditions of selection differ from the present ones he must judge for himself how well the present results ought to extrapolate, and unfortunately the literature does not usually contain much information that will help him to do so.

SUMMARY

The thesis aimed to evaluate selection theory and to compare various techniques of analysing and predicting selection response. It was divided into three parts.

PART ONE: (1) Eighteen generations of selection were carried out for heavier pupae in one population of *Tribolium castaneum* and for lighter pupae in another population. In both cases there were six replications of 64 pupae each, with a separate Control for every replication.

(2) Two methods of analysing the results were compared and the weaknesses of each were assessed. The two methods usually ranked in the same order the relative responses of the three periods of selection, though they often differed considerably in the absolute values of these responses.

(3) The response within each replication was confounded by asymmetry and large fluctuations but the overall result showed genetic gain to decrease with time, slowly to start with then rapidly. There was no clear pattern in the size of fluctuations or trends during each period of the experiment.

(4) Variability between replications was measured each generation by the F-ratio of between-replication variance to within-replication variance. This ratio became significant after only a few generations and increasingly significant as selection continued. Genetic drift was apparently not a major cause of this variability. Variability between replications was also determined for the mean response, standard deviation, and regression coefficient of response over each six-generation period. It was fairly low for the first two parameters but high for the third.

PART TWO: (5) These same populations were studied for predictability of response. A diallel analysis was carried out at the start and end of selection and estimates

of heritability value were made from halfsib, fullsib, parent-offspring regression, and linear heritability analyses. Only the last of these was equipped to predict the asymmetrical response and it did so with fair success. Most of the other estimates were greatly inflated by a maternal effect and by non-additive genetic variance, though less so in the base population.

(6) When the first-generation response was used as a predictor it showed only a slight relationship with the mean, fluctuations and trends of response of the period following. Its discrepancy with the mean response was studied more closely and could be greatly reduced by pooling the data from all replications before attempting to predict. This suggested that with bigger populations such a prediction system might be more useful than is the traditional method of genetically analysing a base population.

(7) The first-generation response of each period was a reasonable predictor of the total subsequent response.

PART THREE: (8) In a second experiment selection for heavier pupae was made at three intensities and for twenty generations. At first the Strong treatment gave the fastest response followed by Moderate then Weak. After about 7 generations Moderate took the lead and held it till about Generation 17. For the remainder of selection Weak gave the heaviest populations, and thus the original order was reversed. These mean weight differences between the three treatments were not significant over fifteen or twenty generations but heritability differences usually were, such realised heritabilities being lowest in the Strong treatment and highest in the Weak one.

(9) For each treatment the response was analysed over five, ten, fifteen and twenty generations. In every case the fluctuation between generations was again measured as standard deviation and the trend over generations was again measured from regression coefficient of response. There was no clear

pattern between periods for standard deviation but Strong always showed the greatest fluctuations and Weak the least. In all three treatments the regression coefficient was always negative but its value came closer to zero with time. This indicated that the rate of response was continually falling off and most rapidly so at the start of selection, but there were no systematic differences between the treatments in this rate of decrease.

(10) Variability between replications was again large for regression coefficient and fairly low for mean response and standard deviation. The F-ratio of between-replication variance to within-replication variance was calculated each generation and it increased in all three treatments but most rapidly in Strong and most slowly in Weak.

(11) Prediction of heritability by variance analysis was measured at Generations 0, 11 and 20. The experimental structure allowed only fullsib and parent-offspring regression analysis and neither agreed consistently with realised heritability for the corresponding period. Evidence again suggested that maternalism and non-additive genetic variance were the main causes of discrepancy.

(12) First-generation response again showed a positive correlation with the mean response of any following period but the correlation was too low to be useful for prediction. The discrepancy was not improved by pooling the data from all replications but there was evidence to show that these results were impaired by a faulty Control.

APPENDIX ONE: CORRESPONDENCE

As part of this thesis programme correspondence was made with the following persons, on points of genetic theory or laboratory technique.

R.W. Allard, Agronomy Department, University of California, U.S.A.

A.E. Bell, Population Genetics Institute, Purdue University, U.S.A.

C.C. Cockerham, Institute of Statistics, University of North Carolina, U.S.A.

A.S. Fraser, Genetics Department, University of California, U.S.A.

B. Griffing, Division of Plant Industry, C.S.I.R.O., Australia.

O. Kempthorne, Statistical Laboratory, Iowa State University, U.S.A.

M. Kimura, National Institute of Genetics, Japan.

K. Kojima, Genetics Department, University of North Carolina, U.S.A.

B. Latter, Division of Plant Industry, C.S.I.R.O., Australia.

I.M. Lerner, University of California, Berkeley, U.S.A.

R.C. Lewontin, Zoology Department, University of Chicago, U.S.A.

A. Robertson, Institute of Animal Genetics, Edinburgh, Scotland.

F.W. Robertson, Institute of Animal Genetics, Edinburgh, Scotland.

B. Sheldon, Division of Animal Genetics, C.S.I.R.O., Australia.

A. Sokoloff, Genetics Department, University of California, U.S.A.

and the managers of breweries at Auckland, Hamilton and Palmerston North.

ACKNOWLEDGMENTS

This thesis was supervised by Professor A.L. Rae, who is head of the Sheep Husbandry Department at Massey University. The author expresses very grateful thanks indeed to Professor Rae for his encouraging and meticulous guidance.

Appreciation is extended also to his co-supervisors, Dr F. Cockrem and Mr R. Patchell, and to Mr A.C. Glenday (Applied Maths Division, D.S.I.R.) for their interest and help.

Finally the author acknowledges with gratitude the support and leave of absence given him by the D.S.I.R. authorities, which made this thesis possible.

REFERENCES

Abplanalp, H. 1961 Linear heritability estimates. Gen. Res. 2:439-448

Anderson, R.L. 1942 Distribution of the serial correlation coefficient. Ann. Math. Stat. 13:1-13

Anonymous, 1960 The usefulness of Tribolium as a laboratory animal. Tribolium Information Bulletin 3:3-13

Bailey, N.T.J. 1959 Statistical methods in biology. England.

Bartlett, A.C. 1962 Changes in quantitative traits under selection and irradiation. Diss. Abstr. 23:3995

Bhat, P.N. 1965 Genetical investigation of prediction of asymmetrical response to two-way selection in Tribolium castaneum. Diss. Abstr. 26:1871

Blair, P.V. 1962 Quantitative consequences of mating systems and selection methods. Diss. Abstr. 22:2127-2128

Bray, D.F., Bell, A.E., and King, S.C. 1962 The importance of genotype by environmental interaction with reference to control populations. Gen. Res. 3:282-302

Brown, W.P. and Bell, A.E. 1961 Genetic analysis of a "plateaued" population of Drosophila melanogaster. Genetics 46:407-426

Cavalli, L.L. 1950 The analysis of selection curves. Biometrics 6:208-220

Chapman, A.B. 1946 Genetic and non-genetic sources of variation in the weight response of the immature rat ovary to a gonadotrophic hormone. Genetics 31:494-507

Chung, C.S. and Chapman, A.B. 1958 Comparisons of the predicted with actual gains from selection of parents of inbred progeny of rats. Genetics 43:594-600

Clarke, J.M., Smith, J.M., and Sondhi, K.C. 1961 Asymmetrical response to selection for rate of development in Drosophila subobscura. Gen. Res. 2:70-81

Clayton, G.A., Morris, J.A., and Robertson, A. 1957 An experimental check on quantitative genetical theory I: Short-term effects of selection. J. Gen. 55:131-151

Clayton, G.A. and Robertson, A. 1957 An experimental check on quantitative genetical theory II: Long-term effects of selection. Ibid 55:152-170

Cochran, W.G. and Cox, G.M. 1950 Experimental designs. New York.

Cockerham, C.C. 1954 An extension of the concept of partitioning hereditary variance for analysis of covariance among relatives when epistasis is present. Genetics 39:859-882

Cockerham, C.C. 1956 Analysis of quantitative gene action. Brookhaven Symposia in Biology No.9:53-68

Comstock, R.E. and Robinson, H.F. 1948 The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics 4:254-266

Comstock, R.E. and Robinson, H.F. 1952 Estimation of average dominance of genes. Heterosis 494-516

Cooke, P., Jones, R.M., Mather, K., Bonsall, G.W. and Nelder, J.A. 1962 Estimating the components of continuous variation. Heredity 17 I: Statistical 115-133
II: Genetical 211-236

Craig, J.V. and Chapman, A.B. 1953 Experimental test of predictions of inbred line performance in crosses. J. Anim. Sc. 12:124-139

Crow, J.F. 1955 General theory of population genetics : synthesis. Coldspring Harbor Symposia 20:54-59

Crow, J.F. and Kimura, M. 1955 Some genetic problems in natural populations. Proc. Third Berk. Symp. on Math. Stat. Prob. 4:1-12 (not seen)

Crow, J.F. and Morton, N.E. 1955 Measurement of gene frequency drift in small populations. Evolution 9:202-214

Dawson, P.S. 1964 The use of assortative mating for heritability estimation. Genetics 49:991-994

Dawson, P.S. 1965 Estimation of components of phenotypic variance for developmental rate in Tribolium. Heredity 20:403-417

DeFries, J.C. and Touchberry, R.W. 1961 The variability of response to selection. I: Interline and intraline variability in a population of Drosophila affinis selected for body weight. Genetics 46:1519-1530

Dickerson, G.E. 1955 Genetic slippage in response to selection for multiple objectives. Coldspring Harbor Symposia 20:213-224

Dickerson, G.E. 1959 Techniques for research in quantitative animal genetics. Amer. Soc. Anim. Prod. publ. 57-105

Dickerson, G.E. 1963 Experimental evaluation of selection theory in poultry. Pro. XIth Int. Congr. Gen. 2:747-761

Falconer, D.S. 1953 Selection for large and small size in mice. J. Gen. 51:470-501

Falconer, D.S. 1954 Validity of the theory of genetic correlation. J. Hered. 45:42-44

Falconer, D.S. 1955 Patterns to responses in selection experiments with mice. Coldspring Harbor Symposia 20:178-196

Falconer, D.S. 1961 Introduction to quantitative genetics. Edinburgh.

Falconer, D.S. 1965 Maternal effects and selection response. Proc. XIth Int. Congr. Gen. 763-774

Falconer, D.S. and King, J.W.B. 1953 A study of selection limits in the mouse. J. Gen. 51:561-580

Fisher, R.A. 1918 The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. (Edinb.) 52:399-433

Fisher, R.A. 1929 The genetical theory of natural selection. New York.

Fisher, R.A. 1941 Average excess and average effect of a gene substitution. Ann. Eugen. 11:53-63

Fisher, R.A. and Yates, F. 1963 Statistical tables for biological, agricultural, and medical research. London.

Fraser, A.S. 1957 Simulation of genetic systems by automatic digital computers
II: Effects of linkage on rates of advance under selection. Aust. J. Biol. Sci. 10:492-499
VI: Epistasis. Ibid 13(1960):150-162

Fraser, A.S. 1962 Simulation of genetic systems. J. Theoret. Biol. 2:329-346

Frey, K.J. and Horner, T. 1955 Comparison of actual and predicted gains in barley selection experiments. Agron. J. 47:186-188

Gardner, C.O. 1963 Estimates of genetic parameters in cross-fertilizing plants and their implications in plant breeding. Symp. on Stat. Gen. and Plant Breeding. NAS-NRC Bull 982:225-252

Gates, C.E., Weber, C.R. and Horner, T.W. 1960 A linkage study of quantitative characters in a soybean cross. Agron. J. 52:45-49

Geiringer, H. 1944 On the probability theory of linkage in Mendelian heredity. Ann. Math. Stat. 15:25-57

Gill, J.L. 1965a Effects of finite size on selection advance in simulated genetic populations. Aust. J. Biol. Sci. 18:599-617

Gill, J.L. 1965b A Monte Carlo evaluation of predicted selection response. Ibid 18:999-1007

Gill, J.L. 1965c Selection and linkage in simulated genetic populations. Ibid 18:1171-1181

Goodale, H.D. 1938 A study of the inheritance of body weight in the albino mouse by selection. J. Hered. 29:101-112

Goodwin, K., Dickerson, G.E. and Lamoreux, W.F. 1960 An experimental design for separating genetic and environmental changes in animal populations under selection. Biometrical Genetics (edited by Kempthorne) 117-138

Graybill, F.A. 1961 An introduction to linear statistical models. New York.

Griffing, B. 1960 Theoretical consequences of truncation selection based on the individual phenotype. Aust. J. Biol. Sci. 13:307-343

Haldane, J.B.S. 1924-7 A mathematical theory of natural and artificial selection. I(1924a): Trans. Cambr. Philos. Soc. 23:19-41 II(1924b): Proc. Cambr. Philos. Soc., Biol. Sci. 1:158-163 III(1926a): Trans. Cambr. Philos. Soc. 23:363-372 IV(1926b): Ibid 23:607-615 V(1927): Ibid 23:838-844

Hanson, W.D. 1963 Heritability. Symp. on Stat. Gen. and Plant Breeding. NAS-NRC Bull 982:125-140

Harris, D.L. 1964 Expected and predicted progress from index selection involving estimates of population parameters. Biometrics 20:46-72

Hayman, B.I. 1955 The description and analysis of gene action and interaction. Coldspring Harbor Symposia 20:79-86

Hayman, B.I. 1960 Maximum likelihood estimation of genetic components of variation. Biometrics 16:369-381

Horner, T.W. 1956 Parent-offspring and full-sib correlations under a parent-offspring mating system. Genetics 41:460-468

Howe, R.W. 1956 The effect of temperature and humidity on the rate of development and mortality of Tribolium castaneum (Herbst.) (Coleoptera, Tenebrionidae). Ann. Appl. Biol. 44:356-368

James, J.W. 1965 Response curves in selection experiments. Heredity 20:57-63

Jones, R.M. 1960 Linkage distributions and epistacy in quantitative inheritance. Heredity 15:153-159

Kempthorne, O. 1954 The correlation between relatives in a random mating population. Proc. Roy. Soc. B 143:103-113

Kempthorne, O. 1955 The correlations between relatives in inbred populations. Genetics 40:168-174

Kempthorne, O. 1957 An introduction to genetic statistics. New York.

Kempthorne, O. 1963 The role of system of mating in the determination of means, variances, and covariances in genetic populations. Symp. on Stat. Gen. and Plant Breeding NAS-NRC Bull 982:21-33

Kempthorne, O. and Curnow, R.N. 1961 The partial diallel cross. Biometrics 17:229-250

Kimura, M. 1958 On the change of population fitness by natural selection. Heredity 12:145-167

Kimura, M. and Crow, J.F. 1963a On the maximum avoidance of inbreeding. Gen. Res. 4:399-415

Kimura, M. and Crow, J.F. 1963b The measurement of effective population number. Evolution 17:279-288

Kojima, K. 1961 Effects of dominance and size of population on response to mass selection. Gen. Res. 2:177-188

Kojima, K. and Kelleher, T. 1963 Selection studies of quantitative traits with laboratory animals. Symp. on Stat. Gen. and Plant Breeding NAS-NRC Bull 982:395-422

Kyle, W.H. and Chapman, A.B. 1953 Experimental check of the effectiveness of selection for a quantitative character. Genetics 38:421-443

Latter, B.D.H. 1960 Natural selection for an intermediate optimum. Aust. J. Biol. Sci. 13:30-35

Long, E.R. 1961 Predicted and actual responses during long-term selection for chemical composition in maize. Euphytica 10:368-378

Lerner, I.M. 1954 Genetic homeostasis. Edinburgh.

Lerner, I.M. 1958 The genetic basis of selection. New York.

Le Roy, H.L. 1960 The interpretation of calculated heritability coefficients with regard to gene and environmental effects as well as to genotype-environmental interactions. Biometrical Genetics (edited by Kempthorne) 107-116

Lewontin, R.C. 1962 Interdeme selection controlling a polymorphism in the house mouse. Amer. Nat. 96:65-78

Lewontin, R.C. 1964 The interaction of selection and linkage
I: General considerations; heterotic models. Genetics 49:49-67
II: Optimum models. Ibid 50:757-782

Lewontin, R.C. and Dunn, L.C. 1960 The evolutionary dynamics of a polymorphism in the house mouse. Genetics 45:705-722

Lewontin, R.C. and Kojima, K. 1960 The evolutionary dynamics of complex polymorphisms. Evolution 14:458-472

Lewontin, R.C. and White, M.J.D. 1960 Interactions between inversion polymorphisms of two chromosome pairs in the grasshopper Moraba scurra. Evolut on 14:116-129

Li, C.C. 1955 Population genetics. Univ. of Chicago Press.

Lush, J.L. 1943 Animal breeding plans. Iowa State College Press.

Lush, J.L. 1948 The genetics of populations. Mimeographed.

Malecot, G. 1948 Les mathematiques de l'heredite. (not seen)

Marien, D. 1958 Selection for developmental rate in Drosophila pseudoobscura. Genetics 43:3-15

Martin, G.A. and Bell, A.E. 1960 An experimental check on the accuracy of prediction of response during selection. Biometrical Genetics (edited by Kempthorne) 178-187

Martin, G.A. and Cockerham, C.C. 1960 High speed selection studies. Ibid 35-45

Mather, K. 1943 Polygenic inheritance and natural selection. Biol. Rev. 18:32-64

Mather, K. 1949a Statistical analysis in biology. London.

Mather, K. 1949b Biometrical genetics. London.

Mather, K. 1955 Response to selection : a synthesis. Coldspring Harbor Symposia 20:158-165

Mather, K. and Harrison, B.J. 1949 The manifold effect of selection. Heredity 3:1-52, 131-162

Matzinger, D.F. 1963 Experimental estimates of genetic parameters and their applications in self-fertilizing plants. Symp. on Stat. Gen. and Plant Breeding NAC-NRC Bull 982:253-279

Mendel, G. 1865 Experiments in plant hybridisation (Transl.) Verh. naturf. Ver. in Brunn, Abhandlungen, iv.

Miller, R.H., Legates, J.E. and Cockerham, C.C. 1963 Estimation of non-additive hereditary variance in traits of mice. Genetics 48:177-188

Mood, A.M. and Graybill, F.A. 1963 Introduction to the theory of statistics. New York.

Nei, M. 1963 Effect of selection on the components of genetic variance. Symp. on Stat. Gen. and Plant Breeding NAS-NRC Bull 982:501-515

Nelder, J.A. 1960 The estimation of variance components in certain types of experiment on quantitative genetics. Biometrical Genetics (edited by Kempthorne) 139-158

Reeve, E.C.R. 1961 A note on non-random mating in progeny tests. Gen. Res. 2:195-203

Reeve, E.C.R. and Robertson, F.W. 1953 Studies in quantitative inheritance II: Analysis of a strain of Drosophila melanogaster selected for long wings. J. Gen. 51:276-316

Rich, E.R. 1956 Egg cannibalism and fecundity in Tribolium. Ecology 37:109-120

Robertson, A. 1955a Prediction equations in quantitative genetics. Biometrics 11:95-98

Robertson, A. 1955b Selection in animals : synthesis. Coldspring Harbor Symposia 20:225-229

Robertson, A. 1956 The effect of selection against extreme deviants based on deviation or on homozygosis. J. Gen. 54:236-248

Robertson, A. 1958 Laboratory breeding experiments and animal improvement. Xth Int. Congr. Gen. 199-207

Robertson, A. 1959 Experimental design in the evaluation of genetic parameters. Biometrics 15:219-226

Robertson, A. 1960 A theory of limits in artificial selection. Proc. Roy. Soc. B 153:234-249

Robertson, A. 1961 Inbreeding in artificial selection programmes. Gen. Res. 2:189-194

Robertson, A. 1966 Artificial selection in plants and animals. Proc. Roy. Soc. B 164:341-349

Robertson, F.W. 1955 Selection response and the properties of genetic variation. Coldspring Harbor Symposia 20:166-177

Robertson, F.W. and Reeve, E.C.R. 1952 Studies in quantitative inheritance I: The effects of selection on wing and thorax length in Drosophila melanogaster. J. Gen. 50:414-448

Robinson, H.F. 1963 Rèsumè : Statistical genetics and plant breeding. NAS-NRC Bull 982:429-438

Schlager, G. 1960 Sperm precedence in the fertilization of eggs in Tribolium castaneum. Ann. Entom. Soc. Amer. 53:557-560

Schnell, F.W. 1963 The covariance between relatives in the presence of linkage. Symp. on Stat. Gen. and Plant Breeding NAS-NRC Bull 982:468-483

Sheldon, B.L. 1963 Studies in artificial selection of quantitative characters I: Selection for abdominal bristles in Drosophila melanogaster. Aust. J. Biol. Sci. 16:490-515
II: Selection for body weight in Drosophila melanogaster. Ibid 16:516-541

Shrode, R.R. 1960 Evidence that mating is random in T. castaneum. Tribolium Information Bulletin 3:15-16

Siegel, P.B. 1962 A double selection experiment for body weight and breast angle at eight weeks of age in chickens. Genetics 47:1313-1319

Snedecor, G.W. 1957 Statistical methods. Iowa.

Solomon, M.E. 1951 Control of humidity with potassium hydroxide, sulphuric acid, or other solutions. Bull. Entom. Res. 42:543-554

Sprague, G.F. and Tatum, L.A. 1942 General vs specific combining ability in single crosses of corn. J. Amer. Soc. Agron. 34:923-932

Tantawy, A.O. 1959 Selection limits with sibmatings in Drosophila melanogaster. Genetics 44:287-296

Tantawy, A.O. and Reeve, E.C.R. 1956 Studies in quantitative inheritance IX: The effects of inbreeding at different rates in Drosophila melanogaster. Z. indukt. Abstamm. -u. Vererblehre 87:648-667

Thoday, J.M. and Boam, T.B. 1961 Regular responses to selection I: Description of responses. Gen. Res. 2:161-176

Tribolium Information Bulletin, 1960 University of California.

Van Vleck, L.D. and Henderson, C.R. 1961 Measurement of genetic trend. J. Dairy Sci. 44:1705-1710

Wong, W.C. 1964 Simplified painting technique for individual identification of beetles. Tribolium Information Bulletin 7:93

Woodworth, C.M., Leng, E.R., and Jugenheimer, R.W. 1952 Fifty generations of selection for protein and oil in corn. Agron. J. 44:60-65

Wright, S. 1920 The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. Proc. Nat. Acad. Sci. 6:320-332

Wright, S. 1921a Systems of mating II: The effects of inbreeding on the genetic composition of a population. Genetics 6:124-143

Wright, S. 1921b IV: The effects of selection. Ibid 6:162-166

Wright, S. 1952 The genetics of quantitative variability. Quantitative Inheritance H.M.S.O. 5-41

Wright, S. 1955 Classification of the factors of evolution. Coldspring Harbor Symposia 20:16-24

Young, S.S.Y. 1966 Computer simulation of directional selection in large populations I: The programme, the additive, and the dominance models. Genetics 53:189-205