CONTAMINATION OF SEED CROPS

I. INSECT POLLINATION*

By A. J. BATEMAN

John Innes Horticultural Institution, Merton

(With Eight Text-figures)

Introduction

It is widely recognized that many agricultural and horticultural crops which are propagated by seed are liable to show a steady deterioration in quality and yield. The seed-growers counteract this tendency to some extent by roguing (i.e. destruction of plants not true to varietal type) and constant re-selection, but they have also endeavoured to remove the causes of this deterioration. One cause which has received much attention is contamination, whether it be due to admixture of foreign seed at harvesting, or admixture of foreign pollen during flowering. The former can be dealt with in a relatively straightforward way, by attending to the effective cleaning of threshing machines, etc., between the handling of different varieties of the same species. Its elimination will also diminish the latter kind of contamination because odd plants of one variety in a field of another will be very efficient suppliers of foreign pollen. It will not, however, completely eliminate contamination by foreign pollen which can be transported over relatively large distances through the agencies of insects or wind, against which we cannot take absolute precautions. Before such precautions can be considered our knowledge of the way in which pollen contamination may arise must be increased.

In spite, however, of the absence of any detailed knowledge of this subject seed-growers have had to devise some method of control. This has usually consisted of the provision of large isolation distances between crops which are likely to cross-pollinate. Distances up to a mile or so are commonly recommended (Haskell, 1943), but though few empirical data have so far been published on the effects of isolation distance, such as are available (Bateman, 1946) are in sharp contrast to the distances recommended in the trade. In the paper referred to above, it has been shown that though seed-growers have apparently used excessive isolation distances, they concentrate their efforts on the prevention of contamination between very distinct varieties of crops, and have overlooked the serious dangers inherent in contamination between relatively similar varieties. This latter type of contamination, though hidden ('cryptic'), may have a larger effect on varietal standards as it can give rise to polygenic variation which cannot be readily rogued out.

It appears then that there is a need for some systematic work to determine under standard conditions what are the effects of isolation distance on various crops. These results should then be applied to all varieties whether the contamination involved were obvious or cryptic, subject to the recognition that isolation requirements of crops vary from one to another and that for the same crop they will vary according to the different purposes (élite, stock or commercial) for which the seed is grown (Bateman, 1946).

Much, however, of the data on contamination so far obtained has been produced under

* Adapted from part of a thesis accepted for the degree of Ph.D. in the University of London.

conditions too unstandardized for them to be of great practical value to the seed-grower. Attention has been focused on the isolating effect of distance whilst the effects of isolation in time, size and arrangement of the plots of the contaminant and contaminated varieties have been largely ignored. As will be seen below, the latter factors are equal to distance in their importance in determining contamination.

Another factor which might be expected to influence contamination in entomophilous plants is the relation between the number of pollinating insects and the supply of nectar. This has been shown by Butler (1943) to affect the method of working of insect pollinators. This is to some extent under the grower's control through adjustment of the total amount of crops in flower at once and of the number of hives in the neighbourhood. Similarly in anemophilous crops wind direction and speed will influence contamination and it can be partly controlled by altering the relative spatial positions of varieties and breaking the wind with walls or trees. The characteristic breeding system for each species, to which the most important contributory factor is the degree of natural self-fertilization (compounded of self-incompatibility, diœcy, monœcy, protandry and even the number of flowers open on a plant at one time) will also influence a crop's liability to contamination. This complex of factors is almost outside the grower's control.

In the work described in the following account an attempt has been made to obtain a more vigorous control than hitherto of contamination-determining factors, and thereby to achieve results more generally applicable to seed-growing practice.

This is not to say that valuable work has not already been undertaken on the effects of isolation distance on contamination. Such is the experiment of Crane & Mather (1943) using two radish varieties (the same as those used in the experiments described below) such that the hybrid could be distinguished from either parent, in which they showed that the relationship between contamination and distance was not a linear one. Where isolation distance was small the rate of decrease of contamination with a given increase in distance was greater than when the isolation distance was larger. They also demonstrated the effect of the mass of plants of one variety growing together on the extent of their contamination by another variety. For in one experiment, in which both varieties were laid out in square blocks with one side in common, the spaces between plants being 9 in., contamination was 0·02 at 15 ft. from the common side, while in an experiment in which a single row of plants of one variety at intervals of up to 24 ft. was strung out from a square plot of the other variety, contamination fell to 0·02 only at a distance of 150 ft.

Currence & Jenkins (1942) studied the effect of distance on contamination in the tomato. Though this crop is self-compatible, many American varieties do not self-pollinate automatically because the style protrudes beyond the cone of anthers. They are thus prone to contamination. Choosing two varieties such that the hybrids were obvious in the first generation, these authors designed an experiment with a central square of one variety and two stringers in opposite directions of the other variety. Contamination in the stringers decreased to a minimum at the maximum isolation distance studied, which was 72 ft.

Balls, Templeton, Brown & Kilain (1929) published results on contamination in cotton. These showed the expected decrease in contamination with increasing isolation distance, but the senior author was undecided whether the contamination was inversely proportional in the distance or to an exponential function of the distance. A marked seasonal fluctuation to the extent of contamination was correlated with seasonal variation in the numbers and kinds of insect visitors. Under Egyptian conditions, hive-bees visited the flowers between epicalyx and corolla thus avoiding pollination. The bee mostly responsible for contamination was a species of solitary bee, and most of the contamination was brought about during a short period when this bee was very active.

There appears to be only one instance in which seed-growing practice has assumed contamination to vary in any definite way with isolation distance. This is contained in the regulations of the Minnesota Crop Improvement Association as reported by Hayes & Immer (1942). It concerns the production of hybrid corn. For a hybridizing plot of size 5 acres, 40 rods (approx. 200 m.) is the minimum isolation distance. This minimum may be reduced, however, if extra border rows of the male parent are planted round the plot in such a way that a reduction in isolation distance of 2½ rods (12·6 m.) can be compensated for by one additional border row. For example, if the isolation distance is 10 rods, twelve border rows must be planted. That the actual effect of isolation distance is not so simply equivalent to border rows is implied by the proviso that border rows may only be substituted for distance when the danger of contamination comes from a variety with the same seed colour as the female parent of the hybrid. Where the intended parents of the hybrid corn had the same seed colour and the F_1 showed no complementary gene action for colour this arrangement would usually involve cryptic contamination.

In the experiments on contamination here presented, the choice of the pair of varieties and which of the pair is to be used as the seed parent of the test progenies, has been made so as to facilitate the early identification of contamination. That is, contamination has always been of an obvious kind, it being assumed that the rules governing cryptic contamination are the same. As interest centres on low frequencies of contamination it is necessary to use large progenies. Earliness in the stage at which contamination is identifiable is important in enabling larger progenies to be grown in the same space and scored in a shorter time. Finally, the crops chosen are those in which contamination is liable to be high, as these are most suitable for work on the external factors concerned with contamination. Such is the case with the following four crops: turnip and radish; both self-incompatible insect-pollinated crops: beet; self-incompatible and generally wind-pollinated, though partially insect-pollinated: and maize; wind-pollinated and self-compatible, but being monœcious and strongly protandrous, a naturally outbreeding crop. The crops beet and maize are dealt with in a later report.

The experiments were designed to study independently the effects on contamination of the distance of the tested variety from contaminating plants and the number of plants of the tested variety growing together. This number is later referred to for simplicity as the mass. The mass does not have a direct effect on the amount of non-contaminant pollen available, as the contribution of pollen from any one plant to the stigmatic surface of another will vary with the distance between them. One can expect, therefore, that the same mass will give variable protection against contamination according to the distribution and density of the plants and the position within the plot of the plant examined for contamination.

It was thought desirable to study the effect of distance over the maximum possible range. For this purpose a series of experiments was arranged, each suitable for the study of the distance effect over a different range. The standard plan consisted of a central plot of the contaminant variety from which extended in various directions arms of varying width and spacing. For such arrangements the problem arises as to what is the zero isolation distance. Since preliminary results suggested that the nearest plants made the biggest pollen contribution, isolation distance was measured from the nearest edge of the contaminant block. At the same time experiments were also designed in which the contaminant was concentrated at a point in the centre of a block of the other variety. Here there could be only one point from which to measure isolation distance. This design also enabled one to vary the density of planting while maintaining the isolation distances and the relative amounts of the two varieties constant.

In an experiment with arms of varying numbers of rows, comparison of the progenies of plants in the same row of the same arm gave the distance effect, and comparison of the progenies of plants in different rows at the same distance gave the mass effect. Such an arrangement is well suited to statistical analysis, permitting independent tests of significance of the various effects.

The dependent variable in these analyses is the proportion of hybrid seedlings to the total progeny, the error variance of a proportion varying with its magnitude according to the general formula, $V_p = pq/n$. As there is a general decrease in this proportion with increasing distance the error variance will always be higher for short distances than for long ones, and the error variance estimated for the whole experiment will be too low for tests of significance at short distances and too high for tests of significance at long distances. But if we transform the frequency of contamination F into an angle ϕ, such that $F = \sin^2 \phi$ (Snedecor, 1946)* the new variable has an error variance independent of its magnitude and depending only on the size of the sample, which shows no trend with distance. The variance of ϕ, measured in degrees, is $820{\cdot}7/n$. Where n, the size of the sample, is not constant, to calculate the theoretical error variance of a number of observations the harmonic mean of the individual sample sizes is used in place of n. The ratio of the sum of squares of error obtained from the analysis of variance to the theoretical error variance gives χ^2 with the same number of degrees of freedom as the error sum of squares. A significantly high χ^2 would show that factors other than sample size were contributory to the error variance.

In the following account, the variety being tested for the contribution of its pollen to the seed of the other variety is referred to as the 'contaminant' and the proportion of progeny produced by hybridization between the two varieties is referred to as the 'contamination' and symbolized by F.

DATA FOR OUT-BREEDING INSECT-POLLINATED CROPS

In this section the two species studied were turnip and radish. The varieties and hybrids used in the experiments are listed below with their diagnostic features and the stage at which detectable.

Crop	Seed parent		Contaminant		Hybrid	
	Name	Appearance	Name	Appearance	Appearance	Stage at which identifiable
Turnip	White Milan	White swollen root	Red-Top Milan	Red top to swollen root	Red top to swollen root	Red pigment shows in young leaves and at leaf bases
Radish	Scarlet Globe	Deep red globular root	Icicle	White long root	Purple longish root	Colour evident in unswollen hypocotyls

* Fisher and Yates *Statistical Tables* (1938) give a less detailed table for angular transformations.

As it has a bearing on the subject of this study it is worthy of note that there appeared to be some contamination in the seeds used for this experiment which were obtained from seedsmen. Seed of Scarlet Globe Radish produced some plants which were apparently identical with the F_1 between that variety and Icicle. In 1945 commercial seed of Scarlet Globe was scored for the frequency of purple rogues of this type, which was found to be 10 out of 1355 or rather less than 1 %. In the author's opinion this is a reasonable amount to expect in commercial seed. However, the frequency of purples only represents a small proportion of the total variability in the variety, as there was a considerable range in the intensity and distribution of the red pigment in the hypocotyl and there was also variation in hypocotyl shape. One might expect that plants of the Icicle variety (which is not a very popular one) would only be responsible for a small proportion of the contamination occurring in Scarlet Globe commercial seed. Amongst Icicle seedlings were found plants with yellow and black hypocotyls.

Seed of White Milan Turnip produced purple-topped plants which on the first occasion, because no roguing had been carried out, were a source of confusion in interpreting the results. Seedlings of White Milan from seed sown in the autumn of 1943 were rogued of all plants showing any anthocyanin in the hypocotyl. This eliminated all purple tops from the experimental plants. A high proportion of green-topped plants remained, though these did not interfere with the experiment.

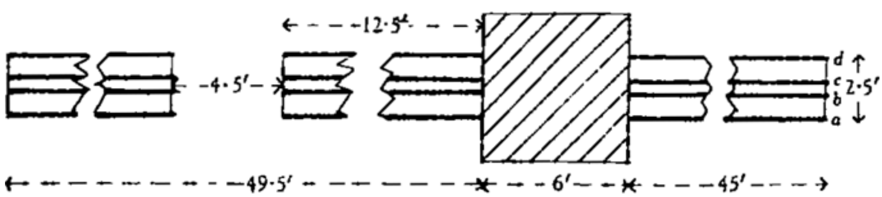

Fig. 1. Plan of Exp. 1. Hatched square represents Red-Top Milan plot from which extend two stringers of White Milan. Heavy black lines represent the four sampled rows.

Experiment 1. This consisted of a central square plot of Red-Top Milan with two arms of White Milan extending from it to east and west. Details of arrangement and dimensions are shown in Fig. 1. Seeds were sown in the autumn of 1942 and the plants planted out in the spring of 1943. Seed was harvested in June 1943. The plants sampled were at various distances from the contaminant plot, in the innermost and outermost lengthwise rows of both arms. The plants in each row were numbered 1–90 beginning from the contaminant plot. As the western arm was interrupted by a path 4 ft. 6 in. wide between plants numbered 25 and 26, each plant in the eastern arm of number greater than 25 had two corresponding plants in the western arm, one of the same number, that is with the same number of plants between it and the contaminant, and one of the same distance from the contaminant, the latter being 4 ft. 6 in. nearer the contaminant than the former. The sampled plants in the eastern arm were numbered 1, 2, 3, 4, 9, 16, 25, 36, 49, 64, 81, 90. Each had a suffix *a, b, c* or *d* according to the row from which it came; *a* and *d* being the outermost rows, and *b* and *c* the innermost. In the western arm plants corresponding to those in the eastern arm were sampled. That means two plants for every one in the eastern arm numbered 36 and higher with the exception of plants 90. The only harvested plants in the western arm corresponding to these were those numbered 90. Each sampled plant had five siliquae harvested from different heights on the main axis of the inflorescence.

1 A and 1 B were the first and second siliquae respectively and 2, 3 and 4 were equally spaced up the remainder of the axis. Siliquae of the same number from different plants did not necessarily coincide in flowering time, but as a whole, siliquae from position 3 had bloomed earlier than siliquae from position 4. As the average number of seeds per siliqua was only about twelve, all siliquae from one plant were pooled in the eastern arm, and all siliquae of the same height from plants of the same number, were pooled in the western arm. In this way, both arms yielded information on the effect of isolation distance on contamination, whilst the eastern arm showed whether there was any effect of row and the western arm whether there was any effect of time of flowering on contamination.

One distance in the western arm (40·5 ft.) has been omitted from the analysis of variance. As will be seen in Table 1, this had an abnormally high value for F, viz. 0·109. During

Table 1

Row no.	Distance from contaminant in ft.	West arm, Position of siliqua: 1 A	1 B	2	3	4	Mean	East arm, Row: *a*	*c*	*d*	*f*	Mean
1	0·5	0·200	0·313	0·564	0·607	0·176	0·372	0·486	0·417	0·071	0·492	0·366
2	1	0·333	0·578	0·406	0·500	0·263	0·416	0·059	0·093	0·073	0·387	0·133
3	1·5	0·460	0·259	0·326	0·391	0·385	0·364	0·340	0·064	0·275	0·382	0·265
4	2	0·231	0·086	0·179	0·364	0·211	0·214	0·333	0·045	0·128	0·406	0·228
9	4·5	0·063	0·094	0·135	0·107	0·188	0·117	0·316	0·300	0·069	0·040	0·181
16	8	0·000	0·019	0·029	0·025	0·037	0·022	0·000	0·029	0·016	0·184	0·057
25	12·5	0·029	0·024	0·027	0·054	0·000	0·027	0·000	0·129	0·000	0·000	0·032
—	18	0·000	0·031	0·000	0·000	0·000	0·006	0·023	0·032	0·000	0·000	0·014
36	—	0·000	0·000	0·000	0·000	0·000	0·000					
—	24·5	0·000	0·107	0·000	0·000	0·000	0·021	0·000	0·000	0·000	0·000	0·000
49	—	0·000	0·000	0·000	0·000	0·000	0·000					
—	32	0·000	0·000	0·000	0·077	0·000	0·015	0·143	0·028	0·000	0·000	0·043
64	—	0·000	0·000	0·000	0·000	0·000	0·000					
—	40·5	0·257	0·094	0·129	0·067	0·000	0·109	0·000	0·000	0·000	0·000	0·000
81	—	0·025	0·000	0·000	0·000	0·000	0·005					
—	45	—	—	—	—	—	—	0·000	0·000	0·000	0·000	0·000
90	—	0·000	0·000	0·000	0·080	0·000	0·016					

Figures for proportion of contamination of White Milan Turnips by Red-Top Milan. Those in the west arm are the result of pooling the seed from corresponding siliquae in the rows represented in the east arm. Those in the east arm are the results of pooling the seed for the five siliquae represented in the west arm for each plant sampled. At distances over 12·5 ft. from the contaminant the west arm has two samples for every one in the east arm, as shown above.

flowering a survey had been made of plants in the White Milan Arm showing abnormally high anthocyanin content. Only one of these suspect plants happened to be sampled for analysis. This was 72 *a*, one of the plants sampled at distance 40·5 ft. It seems justifiable to assume that in this case the high proportion of purple plants was due to one plant being heterozygous for the contaminating gene. Consequently, it was discarded for the analysis.

The analysis of variance in Table 2 leads to the following conclusions: (i) The distance effect is highly significant. This is not surprising. (ii) Position of siliqua (i.e. time of flowering) and (iii) Row, are on the borders of significance. If these effects did prove to have a real foundation they would signify that contamination is highest during the middle of the flowering period, and that it is higher in outermost rows than in innermost rows, which means that it is higher when plants are bordered on only three sides by their own variety than when they are surrounded by them. The latter is borne out by later experiments. As to the former possible effect, since both varieties corresponded very closely in flowering period, significance would imply variation in some external factor such as the relative abundance of the pollinators, which has been shown by Balls *et al.* (1929) to be significant in cotton.

No attempt is made here to explain the results in terms of insect activity. This will be dealt with in a later paper.

Table 2

	Sum of squares	N	Mean square	Variance ratio	Probability
West arm					
Position (i.e. time)	31695·42	4	7923·85	2·32	>0·05
Distance	1604536·35	14	114609·74	33·61	Very small
Error	190984·58	56	3410·44	—	—
Total	1827216·35	74			
East arm					
Distance	745336·23	11	67757·84	7·80	<0·001
Inner and outer rows	28081·69	1	28081·69	3·23	>0·05
Other row effects	16308·37	2	8154·18	—	—
Error	286624·19	33	8685·58	—	—
Total	1076350·48	47			

Analysis of variance of results in Table 1 after they had been submitted to the angular transformation.

The error variance for plants of number over 25 is too great in comparison with contamination for us to judge whether the number of intervening plants between the sampled plant and contaminant plot or the isolation distance is the more important factor in determining the amount of contamination.

Experiment 2. The maximum isolation distance encountered in the above experiment was 49·5 ft. In order to obtain information on the effects on contamination of greater distances, two trials were planned for 1944; one with radish and one with turnip. The designs of both were essentially the same. The contaminant blocks consisted respectively of 121 plants of Icicle Radish and 441 plants of Red-Top Milan Turnip, both being 10 ft. square. The two blocks were side by side. From these in two directions, east and south, plants of Scarlet Globe and White Milan were strung out in straight lines at 20 ft. intervals, up to a distance of 580 ft. in the east stringer and 480 ft. in the south stringer. The experiments were designed in this way to obtain a measurable amount of contamination at all distances, whilst it was hoped that by increasing the general level of contamination the nature of the effect of distance would remain unaltered. All plants in the stringers were harvested.

Considering the radish trial first, thirty capsules, where present, were taken from all parts of the inflorescence. It was considered desirable to keep the factor n, in the expression $V_p = pq/n$ as constant as possible and at the same time as large as possible. So counts were made of the number of seeds harvested in some representative samples. From these it was decided that 150 was the progeny number giving the best combination of constancy and large size. Consequently 150 seeds, when available, were sown from each Scarlet Globe plant. The results are shown in Table 3 and shown graphically in Fig. 2. It is immediately clear that in spite of the much greater distances involved and the different arrangement of the plants the effect of distance on contamination is essentially similar to that in the 1943 turnip trial.

But closer examination shows a new feature in the results. Beyond plant 8 (160 ft.) in either arm, there does not appear to be any further reduction in contamination, as can be seen in Fig. 2. As the error variation is high in comparison with average contamination

Table 3

Plant no.	Distance in ft.	Radish contamination: East stringer	Radish contamination: South stringer.	Turnip contamination both stringers pooled
1	20	0·573	0·642	0·393
2	40	0·434	0·143	0·330
3	60	0·156	0·053	0·258
4	80	0·169	0·106	0·167
5	100	—	0·055	0·170
6	120	0·147	0·049	0·125
7	140	0·084	0·037	0·087
8	160	0·017	0·000	0·030
9	180	0·035	0·024	0·036
10	200	0·000	0·000	0·000
11	220	0·000	0·035	0·086
12	240	0·010	0·017	0·000
13	260	0·000	0·021	0·000
14	280	0·018	0·000	0·000
15	300	0·042	0·007	0·000
16	320	0·000	0·021	0·000
17	340	0·007	0·007	0·000
18	360	0·000	0·000	0·000
19	380	0·008	0·000	0·000
20	400	0·015	0·008	0·000
21	420	0·011	—	0·018
22	440	0·000	0·019	0·105
23	460	0·000	0·025	0·000
24	480	0·000	0·087	0·045
25	500	0·008	—	0·000
26	520	0·021	—	0·000
27	540	—	—	0·000
28	560	0·000	—	0·000
29	580	0·008	—	0·000

Results of long-distance radish trial (two arms given separately), and long-distance turnip trial (result of pooling both arms).

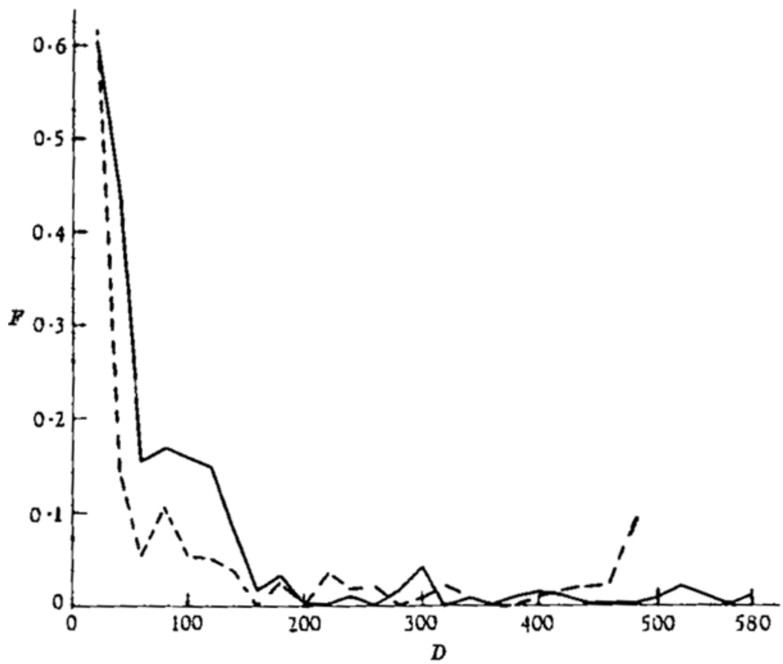

Fig. 2 (Exp. 2). Graph showing effect of distance, *D*, measured in feet, on contamination, *F*. ——, east stringer - - - -, south stringer.

at these distances, it is unwise to draw any conclusions from the appearance of the graph. But we can test the constancy of contamination beyond plant 8 by dividing all the plants concerned into two equal groups, proximal and distal, and summing all the frequencies within each group. We can submit the resulting figures to a χ^2 test.

In the west stringer	Purple	Red	Total	χ^2	Probability
Total for plants 8–18	14	1138	1152	1·40	0·2 to 0·3
Total for plants 19–29	9	1210	1219		
In the south stringer					
Total for plants 8–15	13	794	807	0·23	0·5 to 0·7
Total for plants 16–24	21	1082	1103		

The χ^2 has one degree of freedom. The evidence does not suggest any reduction in contamination from the proximal groups to the distal groups.

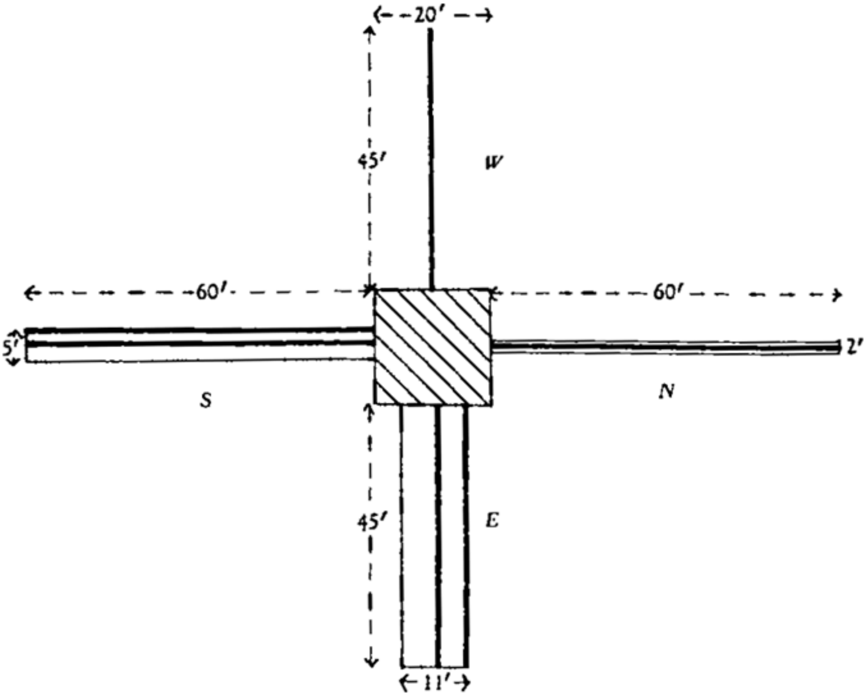

Fig. 3. Plan of Exp. 3. Hatched square represents Icicle plot from which extend four arms of Scarlet Globe. Heavy black lines represent the sampled rows.

In the parallel turnip experiment, depredations by birds forced a premature harvest. Consequently, seed samples were small and of poor germination (average family size 26). Owing to the larger error variance the results of the south and east stringers are pooled and F calculated from the total progenies. These values are shown in Table 3. Though of little value in themselves they serve to confirm the radish results.

Experiment 3. A sub-significant suggestion in the 1943 turnip experiment that the outer rows of a stringer had higher contamination than the inner rows, led to the design and execution of an experiment in 1944, using radish, to test the possible effect of mass of plants on contamination. The lay-out, shown in Fig. 3, consists of a central contaminant block of Icicle with four arms of Scarlet Globe comprising 1, 3, 6 and 12 rows running in W, N, S and E directions respectively. Further data are obtained on the effects of distance, and mass can be studied in two ways; by comparing plants in central rows of each arm, and by comparing plants in innermost (*Ei*, *Si*) and outermost (*Eo*, *So*) rows of the two widest arms. The latter method further provides an opportunity for an estimate of any interaction between mass and distance.

Though the Scarlet Globe sowings were rogued for purple-coloured roots, the process cannot have been complete. For at harvest it was discovered that one plant whose seed was being collected had a purple root (*So* 15). Out of its progeny of 141, 74 were purple, a typical back-cross ratio. Furthermore, as one might expect from the more intense

Table 4

Distance in ft. from contaminant	*W*	*N*	*Si*	*So*	*Ei*	*Eo*
1	0·607	0·586	0·419	0·500	0·278	0·378
2	0·584	0·331	0·154	0·218	0·138	0·250
3	0·250	0·183	0·168	0·321	0·159	0·158
4	0·359	0·275	0·143	0·257	no sample	0·171
5	0·248	0·086	0·132	0·237	0·055	0·089
10	0·068	0·048	0·076	0·074	0·013	0·014
15	0·066	0·007	0·000	*	0·000	0·029
20	0·031	0·029	0·000	0·017	0·031	0·000
25	0·035	0·000	0·005	0·015	*	0·008
30	0·016	0·007	0·000	0·014	0·000	0·000
35	0·014	0·007	0·000	0·000	0·000	0·022
40	0·013	0·007	0·000	0·015	0·000	0·000
45	0·061	0·010	0·014	0·000	0·000	0·000
50	—	0·007	0·000	0·007	—	—
55	—	0·000	0·000	0·009	—	—
60	—	0·000	*	0·007	—	—

Values for contamination in Exp. 3. An * represents a value discarded because there was evidence of contamination brought forward from previous generations (see text).

pigmentation in Scarlet Globe than in the F_1 hybrid, these back-cross purples were more intensely purple than the F_1. Apart from this abnormal family which was discarded from the analysis there were two other plants, *Ei* 25 and *Si* 60 which contained amongst their progeny intensely purple seedlings and showed abnormally high apparent contamination. It was assumed that this also was due to contamination in a previous generation and these progenies were also discarded from the analysis.

Table 5

	Sum of squares	*N*	Mean square	Variance ratio	Probability
		Using series *W*, *N*, *Si* and *Ei*			
Mass	868·128	3	289·376	15·70	<0·001
Distance	7687·326	10	768·733	41·67	Very small
Error	553·465	30	18·449	—	—
Total	9108·919	43			
		Using series *Si*, *So*, *Ei* and *Eo*			
Mass	136·900	1	136·900	12·73	<0·01
Position	85·849	1	85·849	7·98	<0·05
Distance	6520·979	9	724·553	67·38	<0·001
$M \times P$	15·876	1	15·876	1·48	>0·2
$M \times D$	233·970	9	25·997	2·42	0·2–0·05
$P \times D$	101·810	9	11·312	1·05	High
Error	96·775	9	10·753	—	—
Total	7192·159	39			

Analysis of variance for the angular transformation of the result shown in Table 4.

150 seeds, where available, were sown from each plant sampled. These were at distances from the contaminant of 1, 2, 3, 4, 5, 10 ft. and thence every 5 ft. to the end of the arm.

Table 4 shows the proportions of contamination in plants at varying distances and positions in each of the four arms. Fig. 4 shows graphically the results for the central row of each arm. The analyses of variance in Table 5 are carried out on the angular transformation, those distances not represented in all four rows being omitted.

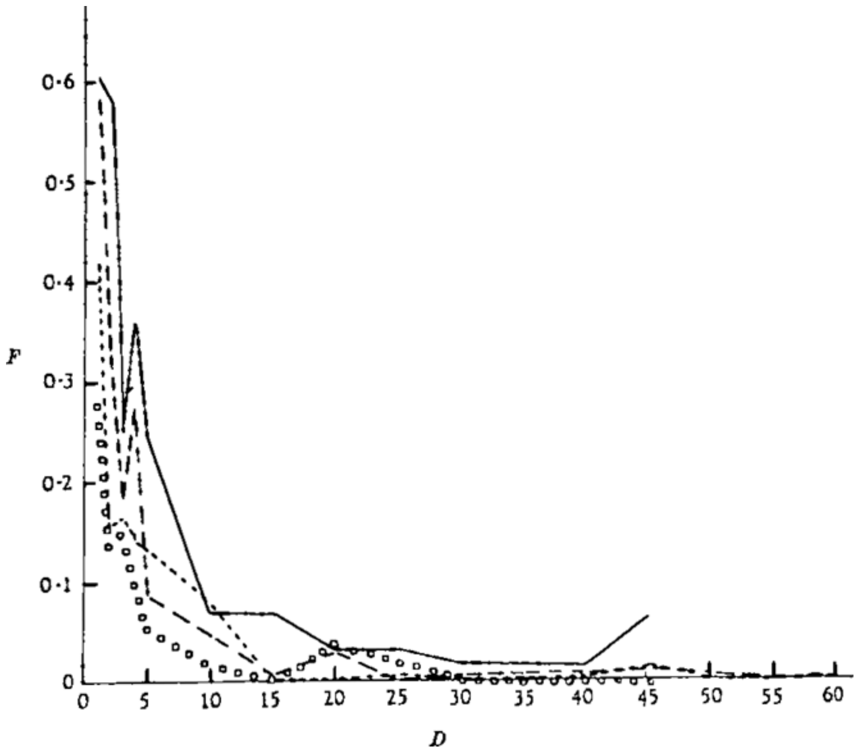

Fig. 4 (Exp. 3). Graph showing effects of distance in feet, D, and width of arm, on contamination, F. ——, one-row arm, W; – – – –, three-row arm, N; - - - -, six-row arm, Si; o o o o, twelve-row arm, Ei.

In the first analysis (for the four centre rows only) the error mean square (18·449) includes any mass-distance interaction there may be. In the second analysis (for two rows of each of the largest arms) this interaction (M.S. = 25·997) is separable from the error mean square (10·753) and does not quite reach the 5 % level of significance.

As the theoretical error variance of ϕ is based solely on the size of the sample, it can be estimated from the data (see introduction). The χ^2 derived by means of this method is a test of whether factors other than sample size make a significant contribution to the error variance estimated from the data.

In the first analysis of variance the theoretical error variance is 8·8115;

$$\chi^2_{(30)} = \frac{553{\cdot}465}{8{\cdot}8115} = 62{\cdot}81;\ P = 0{\cdot}001.$$

In the second analysis the theoretical error variance is 6·3389;

$$\chi^2_{(9)} = \frac{96{\cdot}775}{6{\cdot}3389} = 15{\cdot}27;\ P = 0{\cdot}1 - 0{\cdot}05.$$

From the second analysis we may conclude that sample size can account for the whole error variance. In the first analysis this is not so. As the first error variance includes the mass-distance interaction while the second does not, the mass-distance interaction here being separable and not quite significant, we are justified in assuming that the high χ^2 obtained from the first analysis is due to the mass-distance interaction.

The effects of mass and of the mass-distance interaction can be seen by grouping the families of each arm into two groups, distal and proximal, and calculating the mean value of F for each group. The proximal value is calculated from plants at distances 1, 2, 3, 5 and 10 ft.: the distal value from plants at distances of 20, 30, 35, 40 and 45 ft. The missing distances are those not represented in all rows sampled. The mean values are shown in Table 6 and graphically in Fig. 5. In the latter the logarithm of F has been used so that the relative effects of mass can be compared in the distal and proximal portions.

Table 6

Arm	No. of rows	Proximal	Distal	Proximal/Distal
W	1	0·3514	0·0270	13·01
N	3	0·2468	0·0120	20·57
S	6	0·2299	0·0060	38·32
E	12	0·1532	0·0053	28·91

Exp. 3. The mean contamination in the proximal and distal parts of the four arms (inner and outer rows of *S* and *E* averaged) to show the influence of mass of arm on the relative effect of increased distance from contaminant.

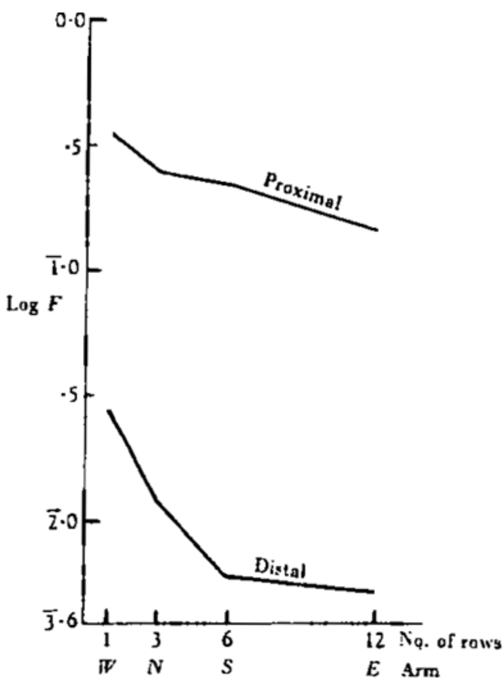

Fig. 5 (Exp. 3). Graph showing effect of isolation distance on the relative decrease in contamination with increasing width of arm. Log F is plotted against row-number for the proximal and distal parts of each arm separately. Proximal is up to 10 ft. Distal is 15 ft. to 45 ft.

It is clear that, generally speaking, increasing the number of rows decreases contamination. In the proximal part of the arms this effect is maintained within the limits of the experiment, from one row to twelve. The relation appears to be described in the graph by a straight line. In other words, the *relative* rate of decrease in contamination produced by the addition of one row remains constant, though of course the *absolute* rate will decrease with increasing numbers of rows. This effect is similar to that of unit increase in isolation distance on absolute contamination, which has been observed in all experiments so far for distances of less than 160 ft.

In the distal part of the arms, however, the relative decrease in contamination per row increase is not constant. It is high where the number of rows is small, higher in fact than the rate in the proximal part, but beyond three rows, increase in row number has a very slight effect and beyond six rows the effect is negligible. We can conclude, that, where the isolation distance is greater than a certain small amount (equivalent to 15 ft. under the conditions of the experiment) and where the mass is above a certain minimum (equivalent to six rows in this experiment), variation in mass has a negligible effect on contamination.

To give a practical application of this conclusion we could state that in the case of a field of a crop being grown for seed at a moderate distance from a contaminant variety, provided the outer rows were discarded, the amount of contamination would be independent of the size of the field.

One possible criticism of the validity of the result of the above experiment is that there may be interference by three of the arms of Scarlet Globe with the results in the fourth. The importance of such interference can be assessed by reference to the experiment.

The maximum interference by one arm with the result of another would be expected in the effect of the twelve-row arm (E) on the one-row arm (W). Let us consider plant $W1$, whose shortest distance from E is 22 ft. At 22 ft. from the contaminant plot, arm W registers a contamination of about 0·03. The contaminant plot has a frontage of twenty-one plants and a total of 441 plants. The arm E has a frontage of twelve plants and a total of 420 plants. We can assume, therefore, that 0·03 is a high estimate of the proportion of seed set on plant $W1$ due to pollen from arm E. By removing arm E we should expect the contamination of plant $W1$ to change from 0·607 as observed to $\frac{0{\cdot}607}{1-0{\cdot}03}=0{\cdot}626$. At 41 ft. from the contaminant, contamination in arm W is about 0·02. By a similar argument removal of arm E would increase the contamination of plant $W20$ from 0·031 to $\frac{0{\cdot}031}{1-0{\cdot}02}=0{\cdot}0316$. We conclude that the effect of arm E is to reduce the contamination in arm W by approximately the same small proportion throughout. The effects of arm E on arms N and S and of other arms on one another would in all cases be rather less. Interference between arms is not, therefore, such as to invalidate the conclusions.

Experiment 4. In all the experiments so far described contamination has been due to a large plot of the contaminating variety. This has been necessary to obtain measurable amounts over the distances used. It is difficult with such a design to know from what point to measure the isolation distance. A superficial examination of the results shows that plants in the immediate vicinity of a given plant are the most important in effecting its fertilization. Thus for plants near to a contaminant plot the isolation distance should be measured from a point near to the nearest edge of the plot. At greater distances it does not matter so much from which point the distance is measured. This difficulty in finding the zero isolation distance can be overcome if necessary by using $(D+\delta)$ for the isolation distance where D is the distance from the nearest edge of the contaminant plot and δ is a correction which may be estimated in some cases from the regression of F on $(D+\delta)$.

This is not reliable when the form of the regression is not known, as the estimated value of δ will vary with the regression used.

So in 1945 an experiment was carried out in which contamination was due to what was in effect a point source, i.e. a small number of plants crowded together. At the same time the possible effects of density of planting were investigated. Four densities were used, each represented by two square plots. Each square consisted of Scarlet Globe Radish plants with a number of Icicle plants at the centre proportional to the density of the plot, so that the ratio of Scarlet Globe to Icicle plants was constant. The internal arrangements of the squares at the four densities are shown in Fig. 6. The experiment was in duplicate, the squares being designated as *WA*, *WB*, *WC*, *WD*, in the western series and *EA*, *EB*, *EC* and *ED* in the eastern replication.

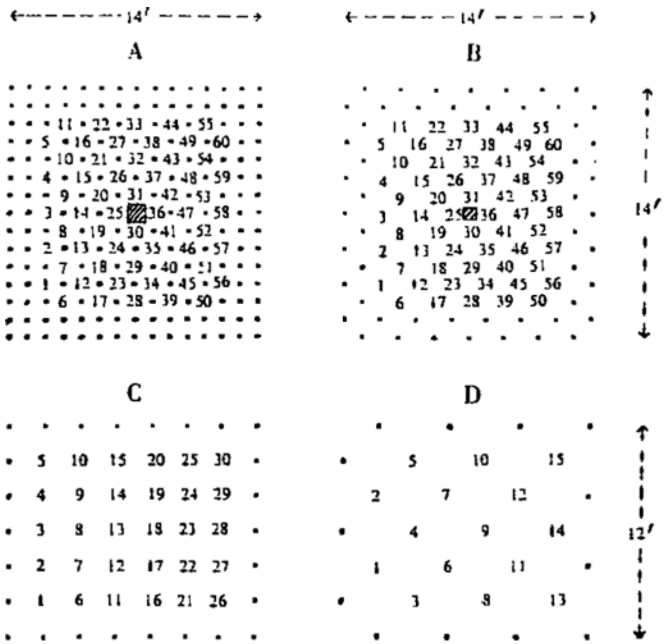

Fig. 6. Plan of Exp. 4. The central hatched area represents 8, 4, 2 and 1 Icicle plants in squares A, B, C, D respectively. All other plants are Scarlet Globe. All sampled plants are numbered, others being represented by a dot each. The squares are not shown in their relative positions. In fact, they were widely separated.

Density was varied in the anticipation that it would affect the extent of distribution of contamination. For, according to Butler (1943), the less the amount of nectar available in a given area for a constant number of working bees the further the bees would travel during foraging. This would lead one to expect that the least dense plots would show the slowest rate of decrease in contamination with increasing distance.

As the Icicle variety is later in flowering than Scarlet Globe, if the whole natural seed set on Scarlet Globe were used to estimate contamination the results would be diluted by a factor depending on the amount of seed set before the Icicle flowered, and to a varying extent for each plant sampled according to the time of onset of its own flowering. To eliminate this source of error, as the group of Icicle plants in each square came into flower all open flowers and capsules from the Scarlet Globe plants in the square were removed. *EA*, *EB*, *EC*, *ED*, *WB* and *WC* were treated in this manner on 30 May. Similarly with

WD on 7 June. In *WA* the Icicle plants were so late that the Scarlet Globe had almost completed flowering, so the contamination was negligible and the plot was valueless for the purpose of the experiment.

There was much depredation of the young capsules by birds. This was particularly marked for squares *EC* and *ED*. As all the outer (later) capsules were removed in these squares the remaining seed would not form a representative sample of the entire seed test.

As there are less plants per distance sampled in the sparser squares, increased error variance in these due to small sample size was avoided by adjusting the size of the sample of seed taken from each plant, so that eight times the amount of seed per plant was harvested from *ED* than from *EA*. There is another source of increased error variance in the sparser squares which was not avoided in the experiment; that due to variation between seed of plants in the same square and at the same isolation distance.

In Table 7 are shown the results of all sampled squares with the plants arranged according to distances. The same results are shown graphically in Fig. 7. The sample progenies

Table 7. *Summarizing contaminations in all the radish density squares* (*Exp.* 4)

Block	Distance in ft.	0·7	1·6	2·1	2·5	2·9	3·5	3·8	4·3	4·5	4·7	4·9	5·1	5·7	6·4
EA	Progeny size	423	900	639	1061	1051	1052	1028	1024	894	816	397	872	849	424
	F	0·111	0·079	0·063	0·055	0·029	0·011	0·015	0·008	0·006	0·013	0·008	0·010	0·006	0·002
	Distance in ft.	1·0		2·2		3·0	3·6		4·1			5·0	5·4		6·4
EB	Progeny size	624		1927		1051	2285		1855			2221	1835		1956
	F	0·223		0·120		0·135	0·047		0·047			0·028	0·026		0·017
EC	Progeny size	705		1070		1033	643		1873			1461	996		660
	F	0·027		0·058		0·105	0·033		0·021			0·021	0·011		0·018
ED	Progeny size	168		275		—	65		8			416	343		46
	F	0·190		0·138		—	0·062		0·250			0·034	0·015		0·022
WB	Progeny size	498		1639		1100	1631		1788			3085	693		1302
	F	0·227		0·103		0·069	0·044		0·026			0·012	0·012		0·011
WD	Progeny size	88		1752		950	464		1254			1813	1463		1235
	F	0·011		0·046		0·080	0·011		0·013			0·009	0·005		0·003
	Distance in ft.	1·4				3·2			4·2			5·1	5·8		7·1
WC	Progeny size	2310				4188			1810			3151	4071		2057
	F	0·119				0·053			0·017			0·016	0·009		0·007

The progeny sizes for each distance are the sums of the progenies of all plants at that distance.

from the outermost rows of squares are not included in this table and figure because in these rows the density is always less than in the rest of the square. Owing to slight errors in placing the plants the distances of sampled plants from the contaminant group are not the same in *EA* and *WC* as in the other squares. These slight differences in planting do not appear to affect the main results.

In Fig. 7 the squares *EA*, *EB*, *WB* and *WC* show a certain regularity in the diminution of contamination with distance, whereas *EC*, *ED* and *WD* are rather irregular. Before drawing any conclusions from these results it is therefore necessary to submit them to an analysis of variance based on the angular transformation.

Table 8 shows the summarized results of this analysis. Beside the error variance calculated from the results is given the theoretical error variance of ϕ from the formula $820{\cdot}7/n$ where n is the harmonic mean of the sample sizes for each square. This theoretical error variance is generally much lower than that calculated from the observations, indicating that there is a real difference between the amounts of contamination occurring in the progeny of different plants at the same isolation distances in the same squares. This could arise through variation in degree of self-incompatibility, variation in the distribution of the seed set with respect to time (this would only be effective in causing variation in contamination if contamination itself varied with time), or variation within the square of varietal mass which would be produced by variation in the floral mass of individual plants.

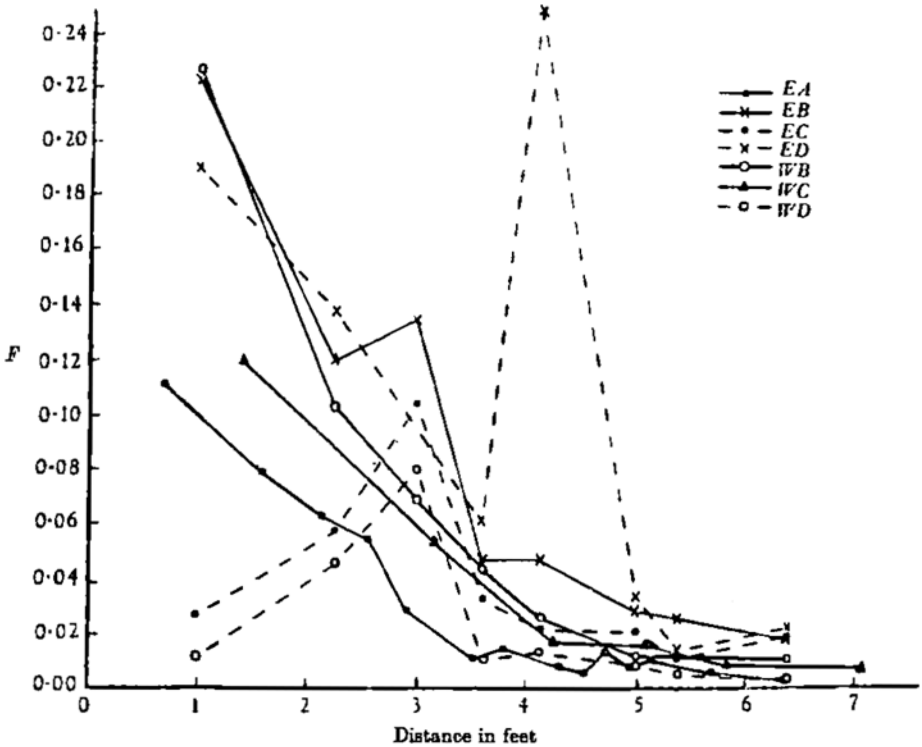

Fig. 7. (Exp. 4). Contamination *F* plotted against distance in ft. for seven squares. Entire lines represent squares with highly significant distance variances; broken lines, squares with less significant distance variances.

Table 8. *Summary of the analyses of variance of the radish density squares based on the angular transformation of* F

	Distance		Error				
Square	*MS*	*N*	*MS*	*N*	$V\phi$	Variance ratio	Probability
EA	168·6699	13	26·7179	84	10·7952	6·31	Very small
EB	267·0905	7	23·0831	50	4·6966	11·57	Very small
EC	39·9888	7	18·0076	15	3·0333	2·22	0·05–0·2
ED	126·4594	6	33·6520	4	22·5545	3·76	0·05–0·2
WB	389·1958	7	39·2327	44	4·5990	9·92	Very small
WC	179·0683	5	10·8679	29	2·0514	16·48	Very small
WD	27·9322	7	1·4312	7	2·0858	19·52	<0·001

The column headed $V\phi$ is the error variance of the angular transformation due to sample size according to the formula $820{\cdot}7/n$ where n is the harmonic mean of the sample sizes of the squares.

The variance ratios of mean square for distance to mean square for error are very high in *EA*, *EB*, *WB* and *WC*. These are therefore reliable sources of data for studying the relation between distance and contamination. In *EC* and *ED* variance ratios are low. In *WD* though high it is based on an abnormally low error mean square, lower even than the theoretical variance calculated from the sample size. If we take the more normal error mean square for *WC* which includes a sample size error variance of equal magnitude the new variance ratio has a probability of 0·05 which is very much less significant than that for *EA*, *EB*, *WB* and *WC*. We are therefore justified in only considering these last four squares for further analysis of the distance effect.

In spite of the great contrast in range of isolation distances (0·7–7 ft. in this experiment as against 20–580 ft. in the long-distance experiment) the essential shapes of the curves relating contamination and distance are the same.

One can also see at a glance that there is considerable variation between squares in the general level of contamination. A figure representing this can be obtained by summing the proportions of contamination of all plants in a square and dividing by the number of Icicle plants at the centre. This gives us the following set of figures:

	Density	*A*	*B*	*C*	*D*
	E	0·3692	1·0169	0·4394	1·0151
Replication	*W*	—	0·7855	0·6404	0·4384

It is seen that though the variation is considerable there is no obvious trend according to the replication or the density.

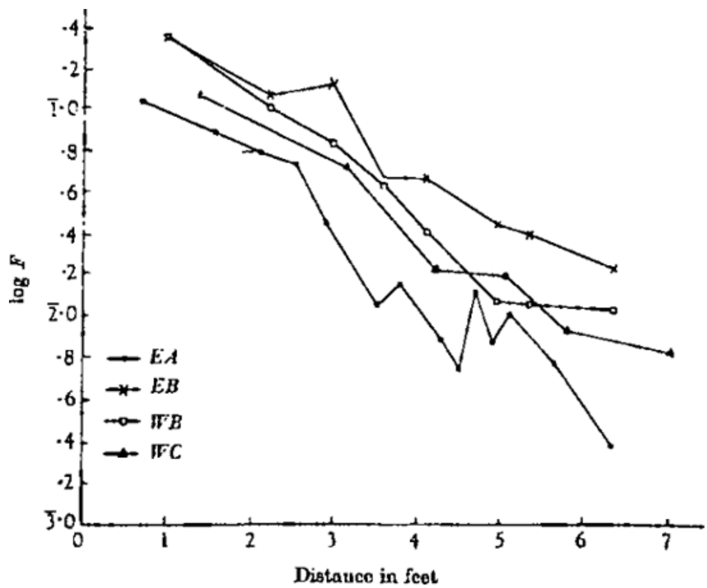

Fig. 8 (Exp. 4). Log *F* plotted against distance in ft. for the four most significant squares.

In examining the effects of density on contamination we must therefore eliminate variation in degree of contamination between squares. This can be done by taking logarithms of the contamination proportions. When plotted on a graph the relation between log contamination and distance will have a slope corresponding to the relative rate of decrease of contamination, i.e. independent of the absolute amounts.

Such curves are shown in Fig. 8. It is clear that all four curves are essentially parallel. This means that density has not been shown to have any effect on the spread of contamination, for an increase in the spread of contamination with decreasing density would cause the slope of *EA* to be greater than that of *EB* and the slope of *WB* greater than that of *WC*. This might at first appear to conflict with Butler's conclusions (1943) regarding the method of action of bees, which led one to expect sparse spacing of flowering plants to cause the bees to forage more widely and produce more widespread contamination. But this would only be so if the number of bees per plot were constant, whereas they were in fact free to distribute themselves as they pleased. The above results suggest that under these conditions the bees tend to distribute themselves in a way (similar to the economic law of supply and demand) such that all foraging bees would have approximately equal returns of nectar or pollen for the same expenditure of energy. In this way the sparser planted squares would have a correspondingly sparser population of pollinators and the relative decrease of contamination with increase in distance would remain constant irrespective of density of planting.

These results were corroborated by an earlier experiment with turnip. Here the error variance was higher and the results therefore were less reliable, but there was again no evidence of any effect of density of planting on spread of contamination.

Summary

1. There are two types of contamination of seed crops, mechanical admixture of seed and cross-pollination. The latter is the more difficult to eradicate.

There are four main variables affecting cross-pollination between varieties: (i) the breeding system of the species; (ii) isolation distance; (iii) varietal mass; (iv) pollinating agent. Of these (iii) and (iv) have been most neglected.

There is little work available entailing a systematic study of these four factors. There is, therefore, a need for such a study of the independent effects of these factors and their interactions.

2. Using two insect-pollinated self-compatible species, radish and turnip, and suitable amounts of plants of the contaminating and contaminated varieties, the effect of distance on isolation is similar over a range of from 0 to 160 ft. At first there is a rapid reduction in contamination with increasing distance, but there is a progressive reduction in the decreases in contamination produced by repeated increases in isolation distance. In the experiment most conducive to contamination this fell from 60 % at 20 ft. to 13 % at 80 ft., but at 140 ft. it was still 6 %. In this same experiment, involving greater isolation distances than the others, on increasing the isolation distance from 160 to 580 ft. there was no evidence of a further decrease in contamination, which remained in the neighbourhood of 1 %. As there are *a priori* reasons for expecting some decrease it must be presumed to be too slight for detection.

3. The mass of a variety (here altered by varying the number of rows from 1 to 12), and the spatial arrangement of the plants, are as important in their effects on contamination as isolation distance. Mass is also similar to isolation distance in its mode of action, in that constant increases in the number of plants of a variety growing together produce progressively smaller decreases in its contamination.

4. When pollinating insects are free to distribute themselves over a crop they do so in such a way that density of planting has no effect on the rate of decrease of contamination with increasing isolation distance.

This work has been carried out under the auspices, and with the financial assistance of the Agricultural Research Council. The author also wishes to express his gratitude to Dr K. Mather for his advice, particularly in regard to the statistical treatment of the work, and to Miss J. Bentley for her technical assistance.

REFERENCES

BALLS, W. L., TEMPLETON, J., BROWN, C. H. & KILAIN, M. (1929). The natural crossing of cotton flowers in Egypt. *Bull.* 89 *of the Egyptian Min. of Agric.*

BATEMAN, A. J. (1946). Genetical aspects of seed-growing. *Nature, Lond.*, **157**, 752.

BUTLER, C. G. (1943). Management of (bee) colonies for pollination. *Ann. Appl. Biol.* **30**, 195.

CRANE, M. B. & MATHER, K. (1943). The natural cross-pollination of crop plants with particular reference to the radish. *Ann. Appl. Biol.* **30**, 301.

CURRENCE, T. M. & JENKINS, J. M. (1942). Natural crossing in tomatoes in relation to distance and direction. *Proc. Am. Soc. Hort. Sci.* **41**, 273.

HASKELL, G. (1943). Spatial isolation of seed crops. *Nature, Lond.*, **151**, 307.

HAYES, H. K. & IMMER, F. R. (1942). Methods of plant breeding (for Regul. of Minnesota Crop Improv. Assn. on 273.)

SNEDECOR, G. W. (1946). *Statistical Methods* (4th ed.). Iowa State Coll. Press, Ames, Iowa.