637.42.002.612:636.5.082.43:519.2

MEDEDELINGEN LANDBOUWHOGESCHOOL
WAGENINGEN • NEDERLAND • 68-20 (1968)

THE COMPARISON OF HYBRIDS ON EGG QUALITY IN A RANDOM SAMPLE TEST BY MEANS OF AN INDEX

De vergelijking van hybriden op ei-kwaliteit in een Random Sample Test met behulp van een index

M. P. F. C. A. TIMMERMANS[1], M. KEULS[2] and M. VAN ALBADA[1]
[1] *Department of Poultry Husbandry,* [2] *Department of Mathematics, Agricultural University, Wageningen, The Netherlands*

(Received 8-VIII-1968)

H. VEENMAN & ZONEN N.V. - WAGENINGEN - 1968

1. INTRODUCTION

Since 1956 commercial hybrids of laying breeds have been tested in the Random Sample Test at Putten (The Netherlands). The purpose of this test is to give information to poultry farmers on the performances of hybrids which are on the market. In this way a more extensive use of the best hybrids is promoted, which benefits the Dutch poultry industry.

The characteristics which are examined concern partly the factors affecting the financial results, such as the number of eggs per hen housed, egg weight, food consumption, etc., partly the egg quality, which is expressed by five characteristics.

The poultry breeders whose entries with regard to the egg quality are among the best and at the same time give financial results above the year average, have up to now received as incentive a financial premium[1] and a certificate[1] for good egg quality.

All the hybrids tested are specific crosses of two or more closed flocks or strains belonging to a same breed or to different breeds.

Because of the system of breeding in closed flocks, which is compulsory in The Netherlands, the hybrids tested are appreciably reproduceable, so that a comparison between the hybrids makes sense. During the past years a total of 22 different breeds or combinations of breeds were included in the test. The principal breeds used in the combinations were White Leghorn, Rhode Island Red, New Hampshire and Australorp. Table 1 presents the breeds and combinations of breeds entered from 1957 to 1966 inclusive, specified according to the principal breeds.

The number of hybrids which every year can be tested in Putten amounts to 75 in duplicate or 50 in triplicate, in two or three pens of 50 hens respectively, divided over three houses. Consequently each entry can be tested with 100 or 150 hens housed. For further details of the testing methods see the annual reports of the Stichting voor het Fokkerijwezen (1957–1966).

The data concerning the egg quality obtained from this Random Sample Test over the years of hatch 1957 to 1964 inclusive have been the starting point for a statistical study aimed at finding out to which extent data obtained from the Random Sample Test are suitable to evaluate differences between entries with regard to the egg quality. It also proved possible to combine egg quality characters into one quality index, which in practice can be used in the comparison of entries.

This report gives a review of the statistical study.

A brief review follows of the contents of the individual paragraphs.

Paragraph 2 deals with the experimental set up of the Testing Station and the egg quality characters considered. In par. 2.1 separate variables ('components') are introduced for the seasonal influences on these characters. Par. 2.2 demon-

[1]Both discontinued recently.

strates with the aid of some analyses of variance which influences are of importance to egg quality. Par. 2.3 considers the correlations, including those between the three components of one character. The correlations between the 'levels' of the characters were generally negligible, in the residual space (table 5) as well as in the entry space (table 6). In par. 2.4 a definite scheme has been drawn up for carrying out the analyses of variance with the sources: years (Y), repeated entries (E) and the interaction Y × E as residual variance, separately for two groups of breeds.

In paragraph 3 and following is discussed how the deviations per entry for each of the quality characters are reduced to terms of loss, viz. in par. 3.1 and in par. 3.2 these reductions are worked out for the shape index and for all quality characters. In par. 3.3 is discussed how the quality characters can be combined in a new scale to one quality index, according to a simplification of Fisher's formula, explained in Appendix II. Finally, in par. 3.4 is discussed how the analysis of variance over three consecutive years, separately carried out for

TABLE 1. Number of entries per year and per breed or combination of breeds

Breed/combination	Years of hatch										
of breeds	57	58	59	60	61	62	63	64	65	66	67
W	7	10	15	26	32	35	46	32	26	20	17
R	19	–	–	2	5	9	13	4	10	8	7
N	4	–	–	–	–	–	–	–	–	–	–
W × R	14	28	29	38	33	21	4	5	2	2	2
W × N	2	3	1	1	1	1	–	–	–	1	–
W × A	1	2	1	2	1	2	1	–	–	–	–
W × (R × N)	–	–	–	2	1	–	–	–	–	–	–
R × W	–	–	–	–	–	–	2	2	4	3	2
N × W	–	–	–	–	–	1	–	2	1	1	1
R × N	–	–	–	–	–	1	2	1	–	2	2
N × R	1	5	3	2	2	1	–	–	–	–	–
A × R	–	–	–	–	–	2	3	2	3	3	3
A × N	–	–	–	–	–	–	–	–	–	–	–
R × (N × R)	–	–	–	–	–	–	1	–	–	–	–
R × (R × N)	–	–	–	–	–	–	–	1	–	–	–
A × (N × R)	–	–	–	–	–	1	1	1	3	1	1
A × (R × N)	–	–	–	2	–	–	2	1	1	3	3
Remaining	2	2	1	–	–	1	–	–	–	6	5
Total	50	50	50	75	75	75	75	51	50	50	43

A = Australorp R = Rhode Island Red
N = New Hampshire W = White Leghorn
W × (R × N) = White Leghorn male × female from Rhode Island Red male and New Hampshire female.

the White Leghorn and for the remaining entries, provides the necessary parameters for drawing up the formula for the egg quality index. In the same paragraph is also given a brief survey of the consistency of the formulas and the repeatability of the index values.

Lastly in paragraph 4 the final procedure in the comparison of hybrids is discussed.

2. Material and statistical study

Every year since 1956, as part of the Random Sample Test, all eggs laid on two consecutive days, always from two pens (of different houses) of every entry during that year, were examined for the egg quality five times a year with intervals of approximately two months. (Only the eggs from the hens reared in 1962 were examined six times with intervals of eight weeks.)

Measurements were made on the shape index (= width × 100/length), specific gravity (by weighing in air and water), albumen quality (Haugh Units), percentage eggs with blood spots and percentage eggs with meat spots.

The average egg weight was not treated as a quality character because:

1. all entries had a satisfactory egg weight
2. the average egg weight (approximately 59 g) can be considered the optimal weight
3. the trade value per kg of eggs is fairly independent of the average egg weight.

The colour of the egg has so far not been judged because in this country the colour of eggs, apart from differences between white and brown, is not expressed in the price.

In practice it is generally assumed that the preference when judging the shape of eggs would lean towards an intermediary shape index. To test a possible aesthetic preference an investigation was set up in conjunction with The Central Institute for Poultry Research at Beekbergen, which will be more extensively reported elsewhere, but some points should be mentioned here.

Twenty series of six different classes of egg weight, consisting of oval eggs only, ellipsoid eggs only, or partly oval and partly ellipsoid eggs, increasing in shape index from 70 to 80 inclusive, were placed twice by 22 experimenters in order of aesthetic shape appearance at intervals of usually a week. This experiment showed that the shape index played only a minor part in the valuation of eggs. Between shape indices 71 and 75 there is little or no difference in valuation; shape index 70 is only slightly less valued then the indices 71–75. An index above 75 decreases the value rapidly, although even eggs with shape index 80 are sometimes still ranked as first.

Although the aesthetic value of the shape index appears to be of little significance, an investigation by Van Tijen (1962) showed that the percentage of damaged eggs during transport increases progressively with a decreasing shape index. This would favour the selection of a high shape index (or changing of packing methods). We will return to this in paragraph 3.

TABLE 2. The seasonal averages for the group of White Leghorn and for the remaining group of hatch 1965

	Oct.	Dec.	Febr.	Apr.	June	Average
White Leghorn: 26 entries						
Shape index	72.5	72.7	72.2	72.4	72.2	72.5
Specific gravity	85.4	83.6	81.4	78.8	76.9	81.2
Haugh Units	81.9	77.3	75.2	72.1	69.2	75.1
Percentage blood spots	2.1	3.2	3.7	3.9	2.3	3.0
Percentage meat spots	0.2	0.3	0.4	0.2	0.2	0.3
Remainder: 24 entries						
Shape index	74.7	73.5	72.6	72.3	72.1	73.1
Specific gravity	81.1	80.9	79.3	77.1	75.3	78.9
Haugh Units	84.1	79.6	77.3	73.5	70.1	77.0
Percentage blood spots	1.2	0.8	1.0	1.3	1.0	1.1
Percentage meat spots	6.6	7.6	9.4	9.6	10.1	8.6

2.1 *The seasonal influence on the quality characters*

Table 2 demonstrates the seasonal influence on the results of hatch 1965 for the different quality characters. In expressing the degree to which the sensitivity for seasonal influences of these entries varies, a linear and a quadratic seasonal component as well as the year average of every five consecutive measurements were calculated for every quality character of every entry.

To this purpose the vector of the five seasonal averages per entry was multiplied by the vectors (1, 1, 1, 1, 1), (—2, —1, 0, 1, 2) and (+2, —1, —2, —1, +2) successively, and the sum of products for each pair of vectors was divided by 5, 10 and 14 respectively.

The specific gravity of the White Leghorn eggs from table 2 may serve as an example for calculation:

the average (in short: *level*):
$[85.4 \times 1 + 83.6 \times 1 + 81.4 \times 1 + 78.8 \times 1 + 76.9 \times 1]/5 = 81.22$

the *linear* component:
$[85.4 \times (-2) + 83.6 \times (-1) + 81.4 \times 0 + 78.8 \times 1 + 76.9 \times 2]/10 = -2.18$

the *quadratic* component:
$[85.4 \times 2 + 83.6 \times (-1) + 81.4 \times (-2) + 78.8 \times (-1) + 76.9 \times 2]/14 = -.043$

Every quality character is thus represented in three measurements, the year average (in short: level), a linear – and a quadratic component.

If we mark the level with 0, the linear and quadratic components with 1 and 2 respectively, we come to the following terminology for the 15 variables of the five quality characters.

	Level	Linear component	Quadratic component
Shape index (Si)	Si_0 (1)	Si_1 (6)	Si_2 (11)
Specific gravity (Sg)	Sg_0 (2)	Sg_1 (7)	Sg_2 (12)
Haugh Units (Hu)	Hu_0 (3)	Hu_1 (8)	Hu_2 (13)
Percentage blood spots (Bs)	Bs_0 (4)	Bs_1 (9)	Bs_2 (14)
Percentage meat spots (Ms)	Ms_0 (5)	Ms_1 (10)	Ms_2 (15)

2.2 *Analysis of variance of the quality characters*

The five quality characters with their year level and the linear and quadratic components produced 15 variables, to which an analysis of variance was applied.

The first analysis of variance concerned the 37 entries which were tested both in 1961 and 1962. These so called 'repeated entries' are characterized by the fact that they consist of the same crosses and that they are marketed under one and the same name. We agree that in future the term 'entry' means entries which have been tested during two or more years.

The experimental results (per pen) from the years 1961–1962 can be classified according to years, houses and entries. In this case an analysis of variance is applied according to the classification:

	Degrees of freedom
Years (Y)	1
Entries (E)	36
Houses (H)	1
Y × E	36
Y × H	1
E × H	36
Residual (Y × E × H)	36
	147

The house effect proved to be negligible for all variables. The entry effect was very great and significant for nearly all 15 variables. The year effect was significant for the levels of the five quality characters and for some of the remaining 10 variables. An analysis for the years 1959–1960 produced an analogous result.

On the basis of these results it was decided to ignore the house effect in future whereas it was considered better to divide the entries into three groups, viz.

White Leghorn
White Leghorn × Rhode Island Red
Remaining hybrids

The attention would thus be directed less towards the large and selfevident differences between the groups of breeds and more towards the smaller differences between the entries of a same homogeneous group.

Tables 3A and 3B present for the 15 variables the results of an analysis of variance of the same data for the year groups 1959–1960 and 1961–1962 respectively, this time divided according to years and entries, with a subdivision of the entries according to breed groups. The analysis of variance follows the scheme:

	Degrees of freedom	
	1959–1960	1961–1962
Year (Y)	1	1
Entries	21	36
breed groups (Bg)	2	2
entries within breed groups (Ew)	19	34
Y × Bg	2	2
Y × Ew	19	34
Residual	44	74
Total (–1)	87	147

The residual variance is composed of the interactions of Y, Bg and Ew influences with 'houses'.

The entries were divided according to breed groups, as follows:

	1959–1960	1961–1962
White Leghorn	7	14
White Leghorn × Rhode Island Red	13	18
Remaining hybrids	2	5

In view of very large differences between the breed groups it was decided in future to make comparisons only between entries within the breed groups White Leghorn and White Leghorn × Rhode Island Red, which are the largest groups. The remaining group was excluded because of the small number of entries.

In the tables 3A and 3B the F-values of the linear and quadratic components of the quality characters in the column 'entries within breed groups' are relatively small compared with the F-values for the 'levels' (Nos 1–5 inclusive). This is a reason to consider in the next procedure mainly the year levels of the quality characters.

The small F-values for the interaction Y × Ew lead to the consideration that interaction effect is negligible. This opens the possibility of treating this interaction as residual variance against which then an entry effect is tested. On the one hand this has the disadvantage of a somewhat limited number of degrees of freedom of this residual variance. According to the tables 3A and 3B this interaction shows also in some cases a small but still significant effect. This naturally reduces the significance of the entry effect when tested against the interaction. However, this is acceptable, as we are mainly interested in those differences between entries which are proof against year differences.

The foregoing gave rise to the choice of a more simple scheme for the analysis

TABLE 3A. Analysis of variance for the years 1959–1960. Mentioned are for 15 variables the F-values[1] for the different effects.

Effect:	Year	Breed group	Entries within Breed group	Year × Breed group	Year × entries within Breed group
Variable	Degrees of freedom				
	1	2	19	2	19
1 Si_0	5.06	118.39	9.35	1.84	0.49
2 Sg_0	1.02	56.52	3.22	2.11	0.34
3 Hu_0	93.75	23.85	12.46	1.04	1.10
4 Bs_0	7.51	14.88	2.58	1.37	1.10
5 Ms_0	13.33	36.94	7.01	2.01	4.53
6 Si_1	7.28	26.76	3.27	0.66	0.54
7 Sg_1	8.70	10.24	2.14	0.81	0.80
8 Hu_1	0.08	13.58	3.04	0.97	3.37
9 Bs_1	0.29	0.74	0.74	0.49	0.93
10 Ms_1	0.00	0.02	0.66	0.43	0.91
11 Si_2	13.17	1.50	1.82	1.05	0.66
12 Sg_2	17.11	0.29	1.03	1.06	0.06
13 Hu_2	0.66	0.30	0.97	0.45	0.22
14 Bs_2	12.30	0.06	0.97	0.57	0.22
15 Ms_2	0.40	0.67	0.76	0.84	0.26
P = .05	4.02	3.21	1.79	3.21	1.79
P = .01	7.13	5.14	2.29	5.14	2.29

[1] The degrees of freedom of the denominator of the variance quotient F is always 44.

of variance with classifications only according to years, entries and interactions, separately for White Leghorn and White Leghorn × Rhode Island Red. This is worked out in table 4 for the year group 1961–1962. The breed groups here are separated.

The F-values for years and entries now use the interaction Ew × Y. According to the significant F-values clear differences between entries of one breed group still emerge from this simpler set up of the analysis.

Moreover, the F-value of the interaction Ew × Y is given in relation to the error variance (= 'remaining interactions' in the original set up of the analysis). Where this F-value exceeds 1, the F-value for years and entries is smaller than the original F-value (in brackets) with this error variance.

2.3 *Correlations between the quality characters*

In order to obtain an impression of the mutual relation of the 15 = 5 × 3 quality characters, the first step was to calculate, in continuation of the above discussed analyses of variance, the correlations between the components of the 15 factors in the residual space (sources of variation and degrees of freedom as in the tables 3A and 3B).

TABLE 3B Analysis of variance for the years 1961–1962. Mentioned are for 15 variables the F-values[1] for the different effects.

Effect:	Year	Breed group	Entries within Breed group	Year × Breed group	Year × entries within Breed group
Variable	Degrees of freedom				
	1	2	34	2	34
1 Si_0	29.55	229.67	13.81	4.48	3.15
2 Sg_0	237.92	203.66	8.49	0.35	1.60
3 Hu_0	305.59	40.52	5.31	2.22	1.62
4 Bs_0	10.06	47.22	4.16	2.87	2.24
5 Ms_0	62.35	514.78	14.78	16.31	3.50
6 Si_1	0.24	85.19	2.85	7.07	2.39
7 Sg_1	18.18	1.20	3.14	8.23	2.20
8 Hu_1	24.90	3.59	1.38	0.91	1.21
9 Bs_1	1.87	4.56	1.18	0.03	0.92
10 Ms_1	0.34	22.20	1.20	0.06	0.59
11 Si_2	0.51	5.34	1.43	0.64	1.06
12 Sg_2	12.14	3.67	1.17	0.03	0.64
13 Hu_2	0.00	0.44	0.68	0.71	0.75
14 Bs_2	12.17	1.55	0.54	2.98	1.69
15 Ms_2	5.09	0.31	1.84	4.11	1.75
$P = .05$	3.98	3.13	1.58	3.13	1.58
$P = .01$	7.03	4.75	1.91	4.75	1.91

[1] The degrees of freedom of the denominator of the variance quotient F are always 74.

These are a measure for the interrelations between the characters, free of the effects of year, breed groups, entries within breed groups and their interactions.

Table 5 shows these correlations, viz. those between the levels in relation to each other (variables 1 to 5 inclusive), and the others as far as they are significant.

We want to consider some (significant) correlations more closely:

a. *specific gravity*

Variable 2 (Sg_0) correlates negatively with 12 (Sg_2). In view of the downward trend of the specific gravity with a negative curvature (see table 2), the conclusion can be drawn that a higher level of specific gravity coincides with a stronger (negative) curvature. This may mean: the entries with the higher average specific gravity show a stronger decline towards the end of the laying period. So more weight might be given to the specific gravity at the beginning of the laying period.

A similar situation is sketched in figure 1.

b. *Haugh Units*

According to the positive correlation of variable 3 (Hu_0) and 8 (Hu_1) hybrids with an average better egg quality show somewhat less decrease in this quality during the laying season (fig. 2).

TABLE 4. Analysis of variance for the years 1961 and 1962 carried out for White Leghorn and White Leghorn male × Rhode Island Red female.
The F-values are calculated in relation to the interaction Y × E_w, the F-values in brackets are calculated in relation to the residual variance.

White Leghorn

Effect space:	Year (Y)		Entries (E_w)		Y × E_w	Total
Degrees of freedom:	1		13		13	35
Variable	F^{1}_{13}	F^{1}_{28}	F^{13}_{13}	F^{13}_{28}	F^{13}_{28}	
Si_0	0.47	(1.00)	4.12	(8.70)	(2.11)	
Sg_0	57.83	(104.09)	2.51	(4.51)	(1.80)	
Hu_0	190.17	(102.69)	8.39	(4.53)	(0.54)	
Bs_0	3.89	(21.33)	1.64	(8.98)	(5.48)	
Ms_0	3.31	(5.66)	1.23	(2.11)	(1.71)	
P = .05	4.67	4.02	2.57	1.90	1.90	
P = .01	9.07	7.12	3.90	2.49	2.49	

White Leghorn × Rhode Island Red

Effect space:	Year (Y)		Entries (E_w)		Y × E_w	Total
Degrees of freedom:	1		17		17	71
Variable	F^{1}_{17}	F^{1}_{36}	F^{17}_{17}	F^{17}_{36}	F^{17}_{36}	
Si_0	8.87	(28.66)	5.72	(18.46)	(3.23)	
Sg_0	99.15	(105.10)	4.62	(4.90)	(1.06)	
Hu_0	106.12	(156.00)	1.87	(2.75)	(1.47)	
Bs_0	1.27	(0.94)	3.84	(2.84)	(0.74)	
Ms_0	18.58	(46.07)	2.48	(6.14)	(2.48)	
P = .05	4.45	3.98	2.28	1.78	1.78	
P = .01	8.40	7.01	3.24	2.25	2.25	

The correlations for Haugh Units suggest that the hybrids could still be more easily distinguished when more weight is given to the beginning or to the end of the laying period. Also the fact is to be taken into account that the observations become less reliable when the laying percentage decreases, whereas the disturbing influence of the temperature makes itself felt on the Haugh Units when the storage period is longer.

Table 6 gives the correlations in the entry space in 1961–1962 for the two breed groups White Leghorn and White Leghorn male × Rhode Island Red female. (The degrees of freedom are like those in table 4.)

TABLE 5. Correlations between egg quality characters, free from year- and entry effects. Except for the levels, only the significant correlation coefficients are mentioned. In italics in both year groups are the correlations which are significant and which bear the same sign (1959-1960 under the diagonal; 1961-1962 above the diagonal)

Characters	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
															1961-1962
1 Si_0	—	+0.07	*+0.40*	−0.01	0.00										
2 Sg_0	+0.14	—	−0.18	−0.15	+0.16		+0.21[1]		−0.22			*−0.28*			
3 Hu_0	*+0.33*	0.00	—	0.00	+0.02	−0.27		*+0.31*							
4 Bs_0	+0.04	+0.03	−0.07	—	0.00				*+0.53*						
5 Ms_0	−0.15	+0.04	−0.09	−0.18	—					*+0.53*					
6 Si_1				+0.35		—	+0.26		−0.22						
7 Sg_1							—								
8 Hu_1			*+0.48*					—							
9 Bs_1				*+0.50*					—						
10 Ms_1					*+0.31*					—					
11 Si_2						+0.30					—				
12 Sg_2		*−0.45*										—			+0.22
13 Hu_2							+0.34	+0.30				+0.48	—		
14 Bs_2			−0.31			+0.31			+0.53					—	
15 Ms_2										+0.54					—
1959-1960															

Significant correlations:	1959–1960	1961–1962
P = .05	0.29	0.23
P = .01	0.38	0.29
Degrees of freedom:	43	73

[1] See paragraph 2.3

TABLE 6. Correlations between the egg quality characters, from differences between entries (1961–1962). Except for the levels (1 to 5 inclusive) only the significant correlation coefficients are mentioned. (Above the diagonal: White Leghorn; under the diagonal: White Leghorn male × Rhode Island Red female).

Characters	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
														White Leghorn	
1 Si_0	—	+0.07	−0.12	−0.39	+0.07	+0.68									
2 Sg_0	−0.31	—	+0.23	+0.55	+0.06		−0.79					+0.54			
3 Hu_0	+0.25	0.00	—	+0.08	+0.67								−0.76		
4 Bs_0	−0.51	+0.38	−0.37	—	−0.24		−0.68		+0.63			+0.58			
5 Ms_0	+0.30	−0.12	+0.09	−0.56	—										
6 Si_1						—							*+0.54*		
7 Sg_1							—	+0.63				−0.62			
8 Hu_1				−0.57				—							
9 Bs_1				+0.75					—						
10 Ms_1										—					
11 Si_2	−0.50										—				
12 Sg_2											−0.48	—			
13 Hu_2													—		
14 Bs_2	−0.48								0.49					—	
15 Ms_2			+0.59			−0.52									—
WL x RIR															

Significant correlations:	WL	WL × RIR
P = .05	0.53	0.47
P = .01	0.66	0.59
Degrees of freedom :	12	16

Fig. 1

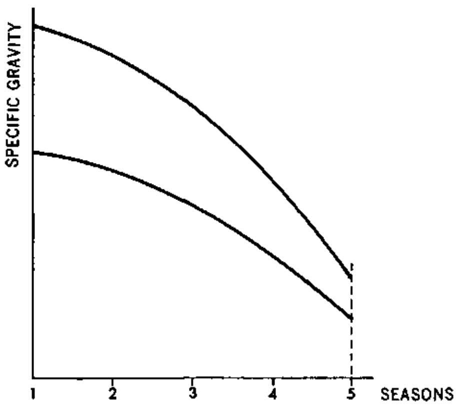
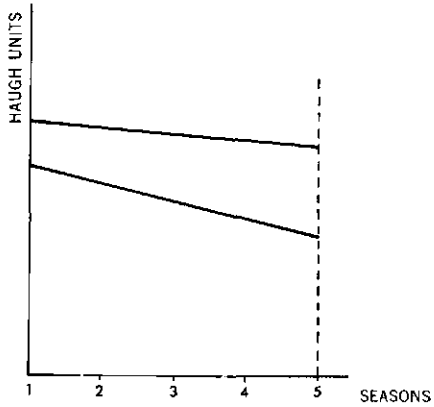


Fig. 2



The correlations in the entry space are as far as the entry component concerns partly due to a genetically determined relation between the quality characters and partly created by differences between the poultry farms where the hatching eggs are gathered.

Significant correlations between the levels of the quality characters are the following. The White Leghorns show a positive correlation between the specific gravity (Sg_0) of the eggs and the percentage blood spots (Bs_0). This correlation is weaker for the White Leghorn × Rhode Island Red. The latter can be explained from the fact that the average specific gravity of the eggs is higher for the White Leghorn than for the White Leghorn × Rhode Island Red. In a previous investigation we found that in general the percentage eggs with blood spots increases progressively with the specific gravity of the eggs. For the year averages this became noticeable at a specific gravity above 1.08; at a lower specific gravity this connection was not noticeable (TIMMERMANS 1964). White Leghorns also show a positive correlation between Haugh Units (Hu_0) and the percentage eggs with meat spots (Ms_0) which is however of little importance, as for this breed the appearance of meat spots is rare. The negative correlation between percentage blood spots and percentage meat spots for White Leghorn × Rhode Island Red may be due to the fact that distinction between these two is often difficult even when the eggs are broken out. For the same reason and because of the comparatively rare occurrence of blood spots in this crossbreed little significance is attached to the correlation between shape index (Si_0 and the percentages eggs with blood spots (Bs_0).

Because the analysis of variance showed that the linear and the quadratic seasonal components of the quality characters contributed little to the differences among the hybrids it seems justified to disregard further the correlations between levels and linear or quadratic components of the quality characters. As we wish to limit ourselves to the levels as they are essential in the comparison of the entries, we want to mention about the correlations of the variables 1 to 5 inclusive (table 6) that these are small in spite of some significancies. We will come back to the positive correlation between specific gravity and percentage blood spots for White Leghorn only in paragraph 3.2.

2.4 *Analysis of variance per group of breeds for the individual characters*

The following procedure for the comparison of the entries was decided on.

a. the year averages (the 'levels') of the quality characters per entry are used only, which produces a very desirable simplification.

b. in the analysis of variance, only year (Y) and entry effects (Ew) are included as main effect; the interaction Y × Ew is taken as the residual variance, against which year and entry effects are tested. Apart from meeting the practical wish for simplification this procedure has the advantage that possible significancies are largely proof against the interaction Y × Ew.

c. the correlations between the levels are neglected; those between specific gravity and percentage blood spots of the White Leghorn group will be discussed in paragraph 3.2.

d. the number of years in the analysis of variance is raised to three, because this enables a sharper distinction between the entries, moreover the number of degrees of freedom of the residual variance will increase. It should be mentioned that not all entries have been repeated in three successive years.

The analysis now includes the entries which have been investigated during at least two of the three years. The classification according to years and entries is here non-orthogonal and the analysis of variance runs according to an iterative calculation procedure (Appendix I).

e. the group White Leghorn × Rhode Island Red is extended with the remaining brown egg laying hybrids.

Table 8 represents as an example the analysis of variance for the White Leghorn and the remaining breeds over the years 1960–1961–1962 and 1961–1962–1963. This table will be further discussed in paragraph 3. We only want to mention the high values FB, which indicate the emergence of clear differences between the entries.

f. it has finally been tried to extend the quality characters with a sixth character, 'the egg quality index', which is really a derivation of the previous five characters. This will be further worked out in paragraph 3. First we discuss in par. 3 a transformation of each of the quality characters making them suitable for addition in such a manner that a quality index is obtained.

3. A COMMON SCALE FOR THE QUALITY CHARACTERS; A QUALITY INDEX

In an attempt to evaluate the characters it was possible to express for each character the deviations from the mean as an extra 'value'. We want to examine the conversion into a common 'value scale' for each character.

3.1 *The shape index in terms of loss*

An investigation by VAN TIJEN (1962) showed that the percentage damaged eggs during transport increases progressively with the decrease of the shape index. This would favour the selection of a high shape index. In view of the fact that there appeared to be no definite preference for a certain shape index (see paragraph 2) it was decided to value the shape index only on the basis of damaged eggs or expected loss during transport.

3.2 *Conversion of the remaining characters in terms of loss*

In accordance with the foregoing an attempt has been made to express all five quality characters concerned as an expected percentage loss.

From the already mentioned investigation by VAN TIJEN (1962) data were available of the percentage damaged eggs and hair cracked eggs which occur during transport among eggs of different classes of specific gravity and shape index (table 7).

TABLE 7. Occurrence of shell damage after transport in relation to specific gravity and shape index.[1]

A. Classification according to specific gravity (December, February and May)

Specific gravity[2]	Number of eggs	Percentages of eggs		
		cracked	broken	cracked and broken
70 and lower	377	31.6	4.5	36.1
71–75	761	14.3	2.4	16.7
76–80	1439	8.5	0.8	9.3
81–85	1709	3.7	0.7	4.4
86–90	767	2.5	0.3	2.8
higher than 90	181	1.1	0.5	1.6
Total and average	5234	8.3	1.2	9.5

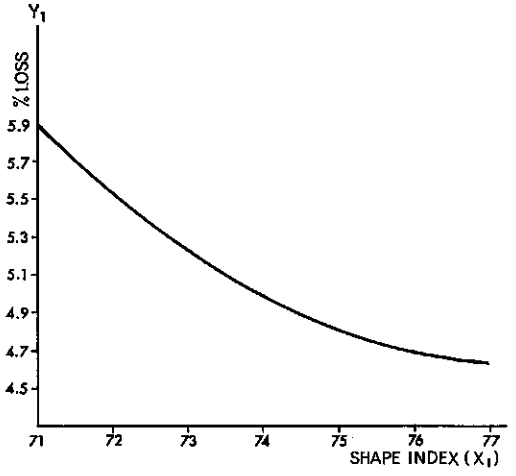
B. Classification according to shape index (February and May)

Shape index	Number of eggs	Percentages of eggs		
		cracked	broken	cracked and broken
Lower than 70	696	12.8	1.3	14.1
70–72	1313	8.8	1.5	10.3
73–75	1412	7.4	1.3	8.7
76 and higher	739	6.5	1.4	7.9
Total and average	4160	8.6	1.4	10.0

[1] From: VAN TIJEN, W. P. (1962), Veeteelt en Zuivelberichten, Vol. 5, no. 2, p. 84.
[2] Factually (SG–1) × 1000.

By counting broken eggs as a complete egg loss and hair cracked eggs as a half egg loss, the (expected) percentages loss (Y_2 and Y_1) could be approximated, on the basis of these data, by quadratic functions of the specific gravity (X_2) or the shape index (X_1) respectively, in which the percentage loss increases progressively with decreasing specific gravity or with decreasing shape index (figures 3 and 4). These formulas (1) and (2) are used in further calculations.

Fig. 3

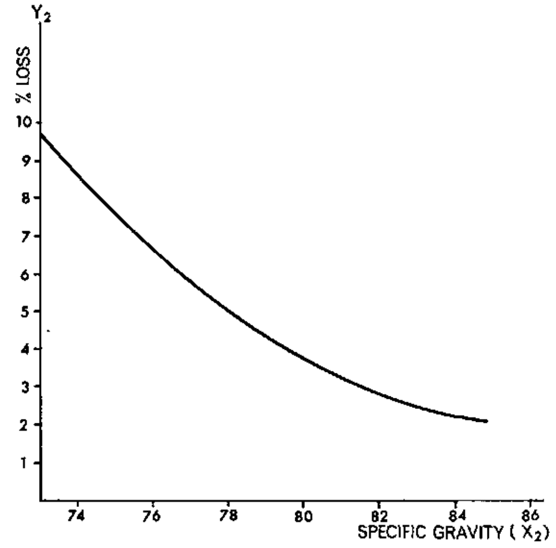


It should be taken into account that the formulas are obtained from the class averages (see table 7) calculated by VAN TIJEN (1962), but that they are applied to the averages of the entries from the Random Sample Test. They would therefore be capable of improvement if more class averages or individual data were available. They are bound to be the same for both breed groups, because no data per breed group were available. They may therefore be influenced by frequencies of breed groups that happened to be present in 1962 as far as the regressions differ from one breed group to another. This point is under investigation.

The conversion into terms of loss has two advantages. In the first place specific gravity and shape index are brought to a common objective basis of valuation. In the second place selection on the basis of these converted values will result in a decrease of the selection pressure according as specific gravity and shape index increase. This will check the increase of the percentage eggs with blood spots which occurs (according to the fairly high correlation coefficients for the characters 2 and 4 in table 6) at a high specific gravity, and at the same time the shape index will not so easily rise above an aesthetically admitted level.

The percentage eggs with blood spots (X_4) can be expressed, though more or less arbitrarily, in figures of loss by counting these as half a loss. If a more severe selection against blood spots is preferred the conversion factor can be set

Fig. 4



higher. For the index to be calculated in this case the factor $\frac{1}{2}$ is used, thus $Y_4 = 0.5\ X_4$.

Haugh Units and percentage eggs with meat spots can not be simply converted into loss figures. On account of a statement by a committee of egg trade experts and persons with practical experience the same importance was attached to improvement of the egg quality (Haugh Units) as to reduction of the percentage blood spots, whereas the same importance was given to the percentage eggs with meat spots as to the shape index.

We thought to express this by making a multiple of the standard deviation σ_{X_3} (= root of the variance in the entry space) of X_3 equivalent to a same multiple of the corresponding standard deviation σ_{Y_4} of Y_4. This term receives a negative sign because a high figure for Haugh Units is considered a good quality and therefore small (or rather negative) loss because of quality faults. Y_4 is the transformed (multiplied by $\frac{1}{2}$) percentage eggs with blood spots.

In a similar way the weighing of the percentage eggs with meat spots X_5 could be carried out by taking a multiple of the standard deviation σ_{X_5} of X_5 equivalent to a same multiple of the corresponding standard deviation σ_{Y_1} of the shape index converted into the loss percentage Y_1.

The original characters X_1 to X_5 inclusive are now expressed in terms of loss Y_1 to Y_5 inclusive, in fact corresponding with the foregoing in the following way:

Shape index: $Y_1 = 0.03X_1^2 - 4.65X_1 + 184.82$ (1)

Specific gravity[1]: $Y_2 = 0.04X_2^2 - 6.95X_2 + 303.86$ (2)

Haugh Units: $Y_3 = -\dfrac{\sigma_{Y_4}}{\sigma_{X_3}} X_3$ (3)

[1] X_2 = (specific gravity − 1) × 1000.

TABLE 8. Analysis of variance

		White Leghorn 60/61/62	White Leghorn 61/62/63	Remainder 60/61/62	Remainder 61/62/63
Shape index (Y_1)	FA	1.49	2.55	3.41	1.45
	FB	4.95	7.29	5.37	7.36
	VB	0.08	0.10	0.12	0.11
	VR	0.02	0.01	0.02	0.02
	f_1	0.80	0.86	0.81	0.86
	$\overline{Y}_1$	'62 5.59	'63 5.49	'62 5.23	'63 5.22
Specific gravity (Y_2)	FA	7.43	12.47	16.72	10.79
	FB	3.64	7.91	10.40	9.92
	VB	0.33	0.72	3.01	4.25
	VR	0.09	0.09	0.29	0.43
	f_2	0.73	0.87	0.90	0.90
	$\overline{Y}_2$	'62 3.14	'63 3.43	'62 4.77	'63 6.40
Haugh Units (X_3)	FA	33.50	31.23	31.51	29.13
	FB	10.91	6.97	4.91	3.72
	VB	6.20	4.99	6.50	6.59
	VR	0.57	0.72	1.32	1.77
	f_3	0.91	0.86	0.80	0.73
	$\overline{X}_3$	'62 75.93	'63 74.68	'62 77.22	'63 76.01
Percentage blood spots (Y_4)	FA	2.21	16.42	0.93	20.85
	FB	1.89	3.04	2.74	4.00
	VB	2.08	2.53	0.84	0.83
	VR	1.10	0.83	0.31	0.21
	f_4	0.47	0.67	0.63	0.75
	$\overline{Y}_4$	'62 3.67	'63 2.11	'62 2.40	'63 0.88
Percentage meat spots (X_5)	FA	2.26	1.21	19.23	8.51
	FB	1.09	1.18	26.19	15.80
	VB	0.33	0.34	48.79	100.87
	VR	0.30	0.29	1.86	6.38
	f_5	0.08	0.15	0.96	0.94
	$\overline{X}_5$	'62 1.19	'63 0.96	'62 10.13	'63 20.28
Degrees of freedom	B	19	24	30	29
	R	22	33	45	34

FA = F-value of the years
FB = F-value of the repeated entries
VB = variance of the repeated entries
VR = residual variance
$f_i = 1 - 1/FB_i$ where $i = 1,\ldots 5$

	Significant F-value for FB: WL P=.05	WL P=.01	Remainder P=.05	Remainder P=.01
60/61/62	2.09	2.86	1.72	2.16
61/62/63	1.86	2.43	1.81	2.32

Percentage blood spots: $Y_4 = 0.5X_4$ (4)

Percentage meat spots: $Y_5 = +\frac{\sigma_{Y_1}}{\sigma_{X_5}} X_5$ (5)

To obtain estimates for σ_{Y_1}, σ_{X_5}, σ_{Y_4} and σ_{X_5} analyses of variance were successively carried out for each three consecutive years and for the (repeated) entries which appeared in two or three of these years. The calculation method, followed in the non-orthogonal design of classes according to years and entries respectively, is described by KUIPER (1952) and CORSTEN (1956). For the evaluation of σ^2, as used in the formulas (3) and (5), the estimate S_1^2 in the 'pure entry' space (see Appendix I) has been taken as a somewhat arbitrary choice. This choice can be considered representative if the corresponding F-values in the 'pure entry' space are high (Table 8).

If they are low this choice does not work too bad, the worst being a low F-value for the Y-variable of reference. For instance the high weight given to Haugh Units in WL (table 9) will be due in part to the low F-value for percentage blood spots in Y_4.

3.3 *Derivation of a formula for the egg quality index*

We assume that the quality of an entry equals 'minus the sum of the expected extra loss percentages' for the five quality characters:

$$\text{quality} = -\sum_{i=1}^{5} (\hat{Y}_i - \bar{\hat{Y}}_i), \quad (6)$$

in which $\bar{\hat{Y}}_i$ represents the average of the expected $\hat{Y}_i$ over the entries. We do not know the expected percentages $\hat{Y}_i$, but we have inaccurate estimators Y_i of $\hat{Y}_i$. For an estimator I of this quality, called quality index, the following formula is chosen:

$$I = -\sum_{i=1}^{5} f_i (Y_i - \bar{Y}_i) \quad (7)$$

in which an estimator $\bar{Y}_i$ of $\bar{\hat{Y}}_i$ is now used, supplemented however with a factor f_i, which will be small when the estimator Y_i is inaccurate. The figures f_i are between 0 and 1. They are estimated from the analyses of variance mentioned in paragraph 3.2 according to (see Appendix II):

$$f_i = 1 - \frac{1}{F_i} \quad (8)$$

in which F_i is the F-value for the repeated entries. In case of F-values < 1 we put $f_i = 0$; in this case the entries do not contribute to the variance of the character and the character concerned does not contribute to the index.

3.4 *Index formulas for two breed groups and the repeatability*

As the entries are now compared within the breed groups only, the most obvious procedure is to use separate formulas for the two breed groups, thus:

a. for the White Leghorn being the only breed group laying white eggs represented in the test, and
b. for all other breeds and crossbreds which lay coloured or slightly coloured eggs. Because of the relatively small number of entries per breed group this group could not stand further divisions.

Apart from differing in egg colour, these two groups also differ in level of shape index, specific gravity, Haugh Units, percentage blood spots and percentage meat spots.

The analysis of variance included the main effects year (Y) and entry within the breed group (Ew), whereas the variance of the remaining interaction Y × Ew was used as denominator in the F-values of the entry and year effects. The analyses of variance, always over periods of three consecutive years concerned the following numbers of entries:

Years of hatch	Number of repeated entries White Leghorn	Number of repeated entries remaining breeds
1959-1960-1961	31	32
1960-1961-1962	20	31
1961-1962-1963	25	30
1962-1963-1964	30	19
1963-1964-1965	31	22
1964-1965-1966	25	21

Table 8 gives the analysis of variance for two of these year groups for both breed groups.

As an example we give the analysis of variance for X_3 in 1961/1962/1963 (White Leghorn).

Source of variation	Degrees of freedom	Sum of squares	Variance	F-value	$f_i = 1 - 1/FB_i$
Repeated entries	24	119.76	4.99(S_1^2)	6.97	0.8564
Years	2	11.06	5.53	31.23	
Residual	33	23.66	0.717	-	

The index formulas, calculated for each three consecutive years, were as follows:

For White Leghorn:

1959/60/61: $I = -0.841Y_1 - 0.735Y_2 + 0.419X_3 - 0.584Y_4 - 0.024X_5 - 21.601$
1960/61/62: $I = -0.798Y_1 - 0.725Y_2 + 0.527X_3 - 0.470Y_4 - 0.040X_5 - 31.466$
1961/62/63: $I = -0.863Y_1 - 0.873Y_2 + 0.608X_3 - 0.671Y_4 - 0.082X_5 - 36.216$
1962/63/64: $I = -0.899Y_1 - 0.917Y_2 + 0.460X_3 - 0.617Y_4 - 0.255X_5 - 24.695$

$$1963/64/65: I = -0.919Y_1 - 0.935Y_2 + 0.417X_3 - 0.646Y_4 - 0.174X_5 - 22.198$$
$$1964/65/66: I = -0.931Y_1 - 0.848Y_2 + 0.364X_3 - 0.801Y_4 - 0.037X_5 - 18.416$$

For the remaining breeds:

$$1959/60/61: I = -0.801Y_1 - 0.829Y_2 + 0.247X_3 - 0.445Y_4 - 0.086X_5 - 8.121$$
$$1960/61/62: I = -0.841Y_1 - 0.904Y_2 + 0.287X_3 - 0.635Y_4 - 0.048X_5 - 11.502$$
$$1961/62/63: I = -0.864Y_1 - 0.899Y_2 + 0.256X_3 - 0.750Y_4 - 0.031X_5 - 7.893$$
$$1962/63/64: I = -0.854Y_1 - 0.875Y_2 + 0.113X_3 - 0.360Y_4 - 0.023X_5 + 1.231$$
$$1963/64/65: I = -0.796Y_1 - 0.900Y_2 + 0.132X_3 - 0.586Y_4 - 0.026X_5 - 1.253$$
$$1964/65/66: I = -0.838Y_1 - 0.954Y_2 + 0.117X_3 - 0.611Y_4 - 0.042X_5 + 0.700$$

In these formulas is:
Y_1 = shape index transformed into percentage loss, according to (1)
Y_2 = specific gravity transformed into percentage loss, according to (2)
X_3 = Haugh Units not transformed
Y_4 = percentage blood spots $\times \frac{1}{2}$ according to (4)
X_5 = percentage meat spots not transformed

These formulas show per breed group a quite acceptable constancy.

Correlations between indices and separate quality characters are given in table 9. This table shows that for White Leghorns the index is determined

TABLE 9. Correlation coefficients of the index for egg quality with the separate quality characters

		Index correlated with:					
		Y_1	Y_2	X_3	Y_4	X_5	
Year	Breed group	Shape index	Specific gravity	Haugh Units	Percentage blood spots	Percentage meat spots	Number of entries
1961	WL	$-.390^1$	+.039	$+.834^2$	$-.651^2$	+.011	32
1962	WL	$-.381^2$	−.119	$+.822^2$	$-.522^2$	+.099	35
1963	WL	$-.636^2$	−.172	$+.865^2$	$-.462^2$	−.052	46
1964	WL	−.097	−.064	$+.811^2$	$-.348^1$	+.207	41
1965	WL	−.276	−.159	$+.596^2$	−.375	+.103	26
1966	WL	−.077	+.099	$+.774^2$	$-.562^2$	−.308	19
1961	remainder	−.060	$-.904^2$	$-.299^1$	$-.340^1$	$-.698^2$	43
1962	remainder	−.043	$-.875^2$	+.077	−.224	$-.530^2$	40
1963	remainder	.000	$-.881^2$	−.098	−.025	$-.475^2$	29
1964	remainder	$+.654^2$	$-.980^2$	$-.409^1$	+.294	$-.600^2$	29
1965	remainder	+.345	$-.965^2$	−.326	−.053	$-.498^1$	24
1966	remainder	+.263	$-.969^2$	−.238	+.064	$-.490^2$	31

[1] $P \leq 0.05$
[2] $P \leq 0.01$

TABLE 10. Some entries of hatch 1964 with their egg quality characters and index value. The index is calculated according to the years 1962-'63-'64

Entry	Shape index X_1	Specific gravity[1] X_2	Haugh Units X_3	Percentage blood spots X_4	Percentage meat spots X_5	Egg quality index
White Leghorn (white eggs)						
56	72.9	77.6	79.4	0.4	0.1	+197
26	72.7	80.1	74.5	1.0	0.7	+ 84
32	71.7	81.4	73.4	2.8	0.9	+ 1
14	73.6	77.3	74.3	1.8	0.2	− 83
9	74.0	78.7	71.1	3.0	0.0	−161
Remainder (slightly coloured eggs)						
42	71.6	81.1	76.9	1.7	4.3	+171
38	73.0	79.1	74.8	1.1	3.5	+106
35	73.6	76.3	79.7	1.9	9.6	− 32
47	73.8	75.5	79.0	2.1	12.0	−102
33	73.8	74.0	76.8	1.2	10.3	−236

[1] Factually (SG−1) × 1000

particularly by Haugh Units and percentage blood spots, for the group of remaining breed groups (short: remainders) by specific gravity and percentage meat spots. These are exactly those characters which need most improvement for the breed groups in question. An example of some results from calculation of the index is given in table 10. The repeatability $R = \sigma^2_b/(\sigma^2 + \sigma_b^2)$ of the quality index per entry and of the separate quality characters is estimated from the formula $S(R) = \frac{F - 1}{F - 1 + b}$ (see Appendix I) in which F is the F-value of the pure entry effect and b the average number of replications (years) per entry calculated as $b = (N - m)/(n - 1)$, in which N = number of observations; m = number of years; n = number of entries.

The results for the years 1961 to 1963 inclusive and 1962 to 1964 inclusive are represented in table 11. The repeatability of the index proves to be somewhat lower for White Leghorn than for the group of the remaining breeds which are (as a group) less homogeneous. In both cases however it is sufficiently high to make the index an acceptable basis of selection. We point out that the repeatability of the index is estimated from the analysis of variance which is also used for the creation of the index formula; this slightly flatters the calculated repeatability of the index.

4. FINAL PROCEDURE FOR COMPARISON OF ENTRIES

Every year the entries in the Random Sample Test are compared. To this purpose those entries are selected which occur at least in two of the last three years. With the aid of an analysis of variance, as previously described, an index

TABLE 11. Repeatability of the quality index and of the separate quality characters for the repeated entries in the year groups 1961-'62-'63, 1962-'63-'64, 1963-'64-'65 and 1964-'65-'66

		Shape index	Specific gravity	Haugh Units	Percentage blood spots	Percentage meat spots	Egg quality index
1961-'62-'63	White Leghorn	0.72	0.74	0.71	0.46	0.07	0.53
	Remainder	0.74	0.80	0.55	0.58	0.87	0.79
1962-'63-'64	White Leghorn	0.78	0.81	0.62	0.39	0.20	0.42
	Remainder	0.70	0.74	0.72	0.19	0.76	0.61
1963-'64-'65	White Leghorn	0.82	0.85	0.62	0.43	0.10	0.43
	Remainder	0.62	0.79	0.76	0.37	0.72	0.65
1964-'65-'66	White Leghorn	0.85	0.71	0.71	0.64	0.01	0.46
	Remainder	0.68	0.90	0.63	0.40	0.93	0.78

formula is calculated with which all entries in the last year are valued in comparison to each other. This procedure is followed separately for the White Leghorn group and for the remaining breeds.

SUMMARY

Egg quality data from the 'Random Sample Test' at Putten from the years of hatch 1957 to 1964 inclusive were subjected to a statistical study with the aim of finding out how the entries can best be compared on the basis of these data. It proved possible to create an egg quality index which enables the distinction of hybrids by one comprehensive figure.

The investigated individual characters X_1 to X_5 inclusive concerned shape index, specific gravity, Haugh Units, percentage blood spots and percentage meat spots respectively. The levels of the five characters, averages over the whole year contained most of the information on differences among the investigated hybrids. Although differences in egg quality between the hybrids, show some season effects these were not so distinct that it seemed worth while to take these into account. The correlations between the five characters in relation to each other were relatively small and have been disregarded in the calculation of a formula for an egg quality index.

As a first result of this investigation a procedure was chosen in which every year analyses of variance are carried out for the five characters, namely for retrospective periods of three years and for the repated entries in three years according to a (non-orthogonal) design of classes: years (Y) and entries (E), in which the year and entry effects are tested against the interaction $Y \times E$. The entries are divided into two breed groups, viz. the 'White Leghorn' and the 'remaining breeds', which are analysed separately. Only entries of a same breed group are compared. Table 10 gives some results.

For the calculation of a quality index the five egg quality characters are transformed into loss percentages. For the shape index X_1 and the specific gravity X_2 the following approximating formulas are obtained on the basis of experimental results by VAN TIJEN (1962):

$Y_1 = 0.03\, X_1^2 - 4.65\, X_1 + 184.82$

$Y_2 = 0.04\, X_2^2 - 6.95\, X_2 + 303.86$, in which $X_2 =$ (specific gravity — 1) $\times 1000$.

The eggs with blood spots were counted as half lost:

$Y_4 = 0.5\, X_4$

For Haugh Units X_3 and percentage meat spots X_5 the loss valuation was obtained indirectly by observing that differences in Haugh Units are equally important as those in percentage blood spots, or those in percentage meat spots equally important as those in shape index, viz.

$Y_3 = -\frac{\sigma_{Y4}}{\sigma_{X3}} X_3$, in which σ^2 [for σ^2 is always taken the estimate S_1^2 of the variance in the pure entry space (perpendicular to years)] is the variance of the entry averages for Y_4 and X_3 respectively;

$Y_5 = +\frac{\sigma_{Y1}}{\sigma_{X5}} X_5$, σ^2 analogous for Y_1 and X_5.

The transformed percentages of loss are added according to the following formula for the egg quality index (I):

$I = -\sum_{i=1}^{5} f_i\, (\hat{Y}_i - \hat{\bar{Y}}_i)$, in which $f_i = 1 - 1/F_i$, $F_i =$ the F-value of the 'pure' entry effect if $F_i > 1$; $f_i = 0$ in case $F_i \leqslant 1$.

The calculated egg quality indices are illustrated in the text as well as in the tables 9, 10 and 11.

In Appendix I and II some aspects of the non-orthogonal design and the index formula are mathematically worked out.

Appendix I

A comment concerning variance-components in a non-orthogonal two-way classification

N. H. Kuiper (1952; § 9) gives the following numerical example of a non-orthogonal two-way classification.

	b_1		b_2	b_3
a_1	314	327	304	285
a_2	329	326	305	306
a_3	269	271	264	

The expectation vector $\hat{y}$ (we write μ) lies by hypothesis in the space D (subspace of R^{11}) spanned by A (row-effects) and B (column-effects):

$$\mu = \mu_D = \mu_{A+B} = a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3 + b_1\beta_1 + b_2\beta_2 + b_3\beta_3,$$

where a_1, a_2, a_3 are the class-characteristic vectors of the subspace A and b_1, b_2, b_3 similarly of B. Summarizing we write:

$$\mu = A\alpha + B\beta\ .$$

where A in matrix-notation represents the set of (column) vectors a_1, a_2, a_3 and α the (column) vector $(\alpha_1, \alpha_2, \alpha_3)'$; (similarly B and β).

A best estimate of μ is the projection y_{A+B} of y onto the space D = A + B, spanned by the vectors of A and B [for a method of calculation see Corsten (1958; p. 36), or in Polish language Nawrovski (1967; p. 253)].

$$y_{A+B} = \begin{bmatrix} 318 & 318 & 303 & 291 \\ 327 & 327 & 312 & 300 \\ 273 & 273 & 258 & \end{bmatrix},\quad y_A = \begin{bmatrix} 307.5 & 307.5 & 307.5 & 307.5 \\ 316.5 & 316.5 & 316.5 & 316.5 \\ 268 & 268 & 268 & \end{bmatrix}$$

$$y_T = y - y_{A+B} = \begin{bmatrix} -4 & 9 & 1 & -6 \\ 2 & -1 & -7 & 6 \\ -4 & -2 & 6 & \end{bmatrix},$$

$$y_{B^0} = y_{A+B} - y_A = \begin{bmatrix} 10.5 & 10.5 & -4.5 & -16.5 \\ 10.5 & 10.5 & -4.5 & -16.5 \\ 5 & 5 & 10 & \end{bmatrix}$$

The analysis of variance for testing the column-effect (B) gives:

$$y^2 = y_A^2 + y_{B^0}^2 + y_T^2$$

Subspace	Dimension	Sum of squares	Variance	F
A (rows)	3	$y_A^2 = 994386$		
B^0 (colums, corrected)	2	$y_{B0}^2 = 1176$	588	12,6
T (rest)	6	$y_T^2 = 280$	46.7	
	11	$y^2 = 995842$		

$P(F_6^2 \geqslant 10.92) = .01$

In the application of this paper the rows (A) correspond with 'years' and the columns (B) with 'entries'. The significant F-value for the subspace B^0 indicates 'significant' differences between the entries.

We now suppose $\underline{\beta}$ to be a stochastic vector (mixed model ANOVA). The vector $\underline{y}$ of observations will have the following distribution [for notation see KUIPER (1959)]:
$\underline{y} \simeq A\alpha + B\underline{\beta} + \underline{t}$, where $\underline{\beta} \simeq \sigma_B \underline{\chi}_n$, $\underline{t} \simeq \sigma\underline{\chi}_N$, $\underline{t}$ and $\underline{\beta}$ independent; N = number of 'plots', n = number of columns, m = number of rows.

First the expectation $E(\underline{y}_{B^0}{}^2) = E(\underline{y}_{A+B} - \underline{y}_A)^2$ will be calculated as a starting-point for the estimation of σ_b (This expectation also determines the power of the F-test for column effects).

Calculation: D = A + B is the subspace spanned by the vectors of A and B, B^0 is the subspace of the vectors in D perpendicular to A.

(N.B. We restrict ourselves to the case that dim $(A + B) = \dim(A) + \dim(B) - 1$, then $d = \dim B^0 = \dim(B) - 1 = n - 1$).

From $\underline{y} \simeq A\alpha + B\underline{\beta} + \underline{t}$ we have $\underline{y}_{B^0} \simeq A_{B^0}\alpha + B_{B^0}\underline{\beta} + \underline{t}_{B^0}$, where A_{B^0} is the set of projections a_{iB^0} $(i = 1,2..m)$ of a_i into B^0, and B_{B^0} similarly. Because of $B^0 \perp A$, A_{B^0} is a set of zerovectors and

$$\underline{y}_{B^0} \simeq B_{B^0}\underline{\beta} + \underline{t}_{B^0}$$

The expectation of the square satisfies:

$$E[\underline{y}_{B^0}{}^2] = E[(B_{B^0}\underline{\beta})^2] + 2E[B_{B^0}\underline{\beta} \cdot \underline{t}_{B^0}] + E[\underline{t}_{B^0}{}^2]$$

Because of independence of $\underline{\beta}$ and $\underline{t}$ the cross-product term equals zero. According to the 'main lemma' of the analysis of variance (KUIPER, 1959, § 5 and 6) the square of the projection of the standard normal N-variate $\underline{\chi}_N$ in R_N onto the subspace B^0 with dimension d is distributed as chi-square with d degrees of freedom:

$$E[\underline{t}_{B^0}{}^2] = E[\{(\sigma\underline{\chi}_N)_{B^0}\}^2] = \sigma^2 E[\underline{\chi}_d{}^2] = d\sigma^2$$

As $\underline{\beta} \simeq \sigma_b\underline{\chi}_n$, we have $E\underline{\beta}_i\underline{\beta}_j = 0,\ i \neq j$, and $= \sigma_b{}^2,\ i = j$. Substitution gives:

$$E[(B_{B^0}\underline{\beta})^2] = E[\Sigma b_{jB^0}\underline{\beta}_j)^2] = \Sigma_j\Sigma_{j'} b_{jB^0} b_{j'B^0}\, E\underline{\beta}_j\underline{\beta}_{j'} = \sigma_b{}^2\Sigma_j b_{jB^0}{}^2$$

For arbitrary $b \in B$ we have $b_{B^0}{}^2 = (b_D - b_A)^2 = (b - b_A)^2 = b^2 - b_A{}^2 = b^2 - \Sigma_i (ba_i)^2/a_ia_i$ as the class characteristic vectors $a_1, .. a_n$ are an orthogonal basis of A. Summarizing we have

$$E[\underline{y}_{B^0}{}^2] = \sigma_b{}^2\Sigma_j[b_j{}^2 - \Sigma_i(b_ja_i)^2/(a_ia_i)] + d\sigma^2$$

We now put: n_{ij} = number of plots in row i and column j, $n_{i.} = \Sigma_j n_{ij}$, $n_{.j} = \Sigma_i n_{ij}$, $N = \Sigma_i n_{i.} = \Sigma_j n_{.j}$. Then it follows that $b_ja_i = n_{ij}$, $a_i{}^2 = n_{i.}$, $b_j{}^2 = n_{.j}$ and by substitution:

$$E\,\underline{y}_{B^0}{}^2 = \sigma_b{}^2\Sigma_j[n_{.j} - \Sigma_i \frac{n_{ij}{}^2}{n_{i.}}] + d\sigma^2 = \sigma_b{}^2[N - \Sigma_i \frac{\Sigma_j n_{ij}{}^2}{n_i}] + d\sigma^2 \qquad (1)$$

In the case that moreover for each combination i j only one or no observation is present we have

$$n_{ij} = 1 \text{ or } 0,\ n_{ij}{}^2 = n_{ij} \text{ and } \Sigma_j n_{ij}{}^2/n_{i.} = \Sigma_j n_{ij}/n_{i.} = 1$$

Summarizing: If n_{ij} takes values 1 or 0 only, we have

$$E(\underline{y}_{A+B} - y_A)^2 = E(\underline{y}_{B_0}{}^2) = d(b\sigma^2{}_b + \sigma^2)$$

$$d = n - 1,\ b = \frac{N - m}{n - 1}$$

m = number of classes of A
n = number of classes of B

The 'repeatability' R is defined as $R = \dfrac{\sigma_b{}^2}{\sigma_b{}^2 + \sigma^2}$

Let us call $s_1{}^2 = y_{B_0}{}^2/(n - 1)$, $s^2 = y_T{}^2/\dim T$, $\dim T = N - n - m + 1$, then $\underline{s}_1{}^2$ is an unbiased estimator of $b\sigma_b{}^2 + \sigma^2$, and $\underline{s}^2$ of σ^2. The variance-quotient $\underline{F} = \underline{s}_1{}^2/\underline{s}^2$ is a 'reasonable' estimator of $b(\sigma_b{}^2/\sigma^2) + 1$, if $\underline{F} > 1$. An estimate of the repeatability R now follows easily:

$$\text{Estimate } \frac{\sigma_b{}^2}{\sigma_b{}^2 + \sigma^2} = \frac{F - 1}{F - 1 + b},\ F > 1 \qquad = 0 \quad F \leqslant 1 \tag{3}$$

The standard deviation of the estimator is difficult to obtain. However, the probability-distribution of the variance-quotient $\underline{F} = \underline{s}_1{}^2/\underline{s}^2$ is related to a central Fishervariate[1] with ν_1 and ν_2 degrees of freedom by

$$\frac{\underline{s}_1{}^2/(b\sigma_b{}^2 + \sigma^2)}{\underline{s}^2/\sigma^2} \simeq \underline{F}^{\nu_1}_{\nu_2} \qquad \nu_1 = n - 1,\quad \nu_2 = \dim T$$

The upper 5% of $\underline{F}^{\nu_1}_{\nu_2}$ is found in an F-table, while a 95% point follows from

$$F^{\nu_1}_{\nu_2}(.95) = 1/F^{\nu_2}_{\nu_1}(.05)$$

With the aid of these tabulated values then follows a 90% confidence interval of R

$$\frac{\underline{F}/F\,.05 - 1}{\underline{F}/F\,.05 - 1 + b} < \frac{\sigma_b{}^2}{\sigma_b{}^2 + \sigma^2} < \frac{\underline{F}/F\,.95 - 1}{\underline{F}/F\,.95 - 1 + b}$$

From this interval the negative portion should be omitted (SCHEFFÉ, 1959).

[1] Although $\underline{s}_1{}^2$ and $\underline{s}^2$ are distributed independently, $\nu_1\underline{s}_1{}^2/\sigma_1{}^2$ follows a chi-square distribution only approximately; therefore the F-distribution is an approximation.

Appendix II

The calculation of a selection-index

$\underline{x}_i$, $i = 1,2\ldots n$ be a set of characters measured on *one entry* with expectations ξ_i. The 'value' of that entry is expressed by $\eta = \sum_{1}^{5} b_i\xi_i$, in which the coefficients b_i are supposed bo be known.

We suppose

$$\underline{x}_i \simeq \xi_i + \underline{u}_i \qquad i = 1,2\ldots n$$
$$E[\underline{u}_i] = 0$$
cov. matrix: $$\{E[\underline{u}_i\underline{u}_j]\} = \frac{1}{r}\,U,$$

where r is the number of replicates, of which the measurement (vector) is an average.

We now consider entries taken *at random* from a universe and with it the stochastic sets $\underline{\xi}_i$, $i = 1,2\ldots n$ and corresponding $\underline{x}_i$, $i = 1,2\ldots n$ (measurements of ξ_i).

We suppose

$$E[\underline{\xi}_i] = 0 \qquad i = 1,2\ldots n$$
cov. matrix: $$\{E[\underline{\xi}_i\underline{\xi}_j]\} = X$$
$$(\underline{u}_1, \underline{u}_2\ldots\underline{u}_n), (\underline{\xi}_1, \underline{\xi}_2\ldots\underline{\xi}_n) \text{ independent.}$$

We define $\underline{y} = \Sigma a_i\underline{x}_i$ to be the *index of the value* of an entry (chosen at random) where the coefficients a_i are to be chosen in such a way, that taken over the population of entries, $\underline{y}$ and $\underline{\eta}$ are correlated maximally. That means the correlation coefficient $\rho = E\,[\underline{y}\underline{\eta}]/\sqrt{E\,[\underline{y}^2]\,.\,E\,[\underline{\eta}^2]}$ is maximized as a function of $a_1,\ldots a_n$.

Calculation of $a = (a_1, a_2\ldots a_n)'$.

In the following a matrix-notation is used because of brevity, thus a is a columnvector, similarly $\underline{u} = (\underline{u}_1, \underline{u}_2\ldots\underline{u}_n)'$. etc.

$$\frac{1}{r}U = E\,[\underline{u}\,\underline{u}'],\ X = E\,[\underline{\xi}\,\underline{\xi}'],\ E\,[\underline{u}\,\underline{\xi}'] = 0$$

Then

$$E\,[\underline{\eta}^2] = E\,[(b'\,\underline{\xi})^2] = b'\,E\,[\underline{\xi}\,\underline{\xi}']\,b = b'Xb$$
$$E\,[\underline{y}\,\underline{\eta}] = a'Xb$$
$$E\,[\underline{y}'\underline{y}] = a'\left(X + \frac{1}{r}U\right)a \overset{\text{p.d.}}{=} a'Wa, \text{ where } W = X + \frac{1}{r}\,U$$

Thus:

$$\rho = \frac{a'Xb}{\sqrt{(b'Xb.a'Wa)}}$$

A non-singular symmetrical matrix W can be written: $W = D'D$, where D'^{-1} exists. Let us now call $a^* = D \cdot a$, $b^* = D'^{-1}Xb$, then:

$$\rho = \frac{a^{*\prime}b^*}{\sqrt{[(a^{*\prime}a^*)\,(b'Xb)]}}$$

Now the length $(a^* \cdot b^*)/\sqrt{(a^* . a^*)}$ of the orthogonal projection of the vector b^* on a^* is a maximum if a^* has the direction of b^*. Therefore ρ is a maximum if $a^* = \lambda b^*$.

By backward substitution, this implies $a = \lambda W^{-1}Xb$
and if we choose $\lambda = 1$:

$$a = W^{-1}Xb, \ \rho = \frac{b'Xa}{b'Xb} \tag{1}$$

Estimates of matrices X and W are obtained by means of an analysis of variance as described in appendix I, but now for a series of characters $i = 1,2,\ldots n$, and applied to a random sample from the universe of entries.

The variances and covariances obtained in the subspace B^0, together give an estimate of the covariance matrix $bX + U$ where B^0 and b are defined as in appendix I. In the subspace T of residuals an estimate is obtained of the covariance matrix U.

In the special case that

1. $r = b$
2. The correlations in the subspace B^0 and T are negligibly small (then U and $W = X + \frac{1}{r} U$ and therefore X are diagonal matrices), the (diagonal-) elements of $W^{-1}X$ are:

$$\left(\frac{\sigma_b^2}{\sigma_b^2 + \frac{\sigma^2}{r}}\right)_i \quad i = 1,2\ldots n$$

Estimates of these elements are: $(1 - 1/\underline{F}_i)$ (if $\underline{F}_i > 1$), as: $\underline{F}_i = (\underline{s}_1^2/\underline{s}^2)_i$ and $(E\,\underline{s}_1^2)_i = (r\sigma_b^2 + \sigma^2)_i$, $E(\underline{s}^2)_i = (\sigma^2)_i$

Summarizing: when $\lambda = 1$ and with the forementioned restrictions we have the following estimators $\underline{a}_i$ of a_i.

$$\underline{a}_i = (1 - 1/\underline{F}_i)b_i, \quad \underline{F}_i > 1 \quad i = 1,2\ldots n$$
$$= 0 \qquad \underline{F}_i \leqslant 1$$

In the general case the technique described is equivalent to that described by FISHER (1954; § 49.2).

REFERENCES

CORSTEN, L. C. A. (1958), Vectors, a tool in statistical regression theory, Thesis. Agric. Univ. Wageningen.

FISHER, R. A. (1954), Statistical methods for research workers. Oliver and Boyd, London.

KUIPER, N. H. (1952), Variantieanalyse, Statistica Neerlandica, 6, 149.

KUIPER, N. H. (1959), Een symbool voor isomoor, Statistica Neerlandica 13, 295.

NAWROCKI, Z. (1967), Teoria i praktyka doswiadczenia rolniczego, Warszawa.

SCHEFFÉ, H. (1959), The analysis of variance. Wiley and Sons, New York.

Stichting voor het Fokkerijwezen bij de Pluimveehouderij, Zeist. Jaarverslagen 1957–1966.

TIMMERMANS, M. P. F. C. A. (1964), Correlaties tussen eikwaliteitsfactoren. De Bedrijfspluimveehouder, 42, nr. 6, p. 201.

TIMMERMANS, M. P. F. C. A., M. VAN ALBADA en M. KEULS (1966), An egg quality index for discrimination between hybrids in a random sample test. Proceedings XIIIth World's Poultry Congress.

VAN TIJEN, W. P. (1962), Het verband tussen enkele factoren van de eikwaliteit en het percentage rammelaars en andere breuk bij normaal vervoer. Veeteelt- en Zuivelberichten, nr. 2, p. 84.