FACTOR ANALYSIS OF PLANT-TYPE VARIABLES RELATED TO YIELD OF DRY BEANS (PHASEOLUS VULGARIS L.)

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

JACQUES CLAREL DENIS

1971

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 00823 9778

L

LIBRARY
Michigan State
University

This is to certify that the

thesis entitled

Factor Analysis of Plant-Type Variables Related to Yield of Dry Beans (*Phaseolus Vulgaris* L.)

presented by

Jacques Clarel Denis

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Crop and Soil Sciences Department

M. W. Adams
Major professor

Date November 12, 1971

O-7639

BINDING BY
HOAG & SONS' BOOK BINDERY INC.
LIBRARY BINDERS
SPRINGPORT, MICHIGAN

289

0430202
AUG 22 2002

ABSTRACT

FACTOR ANALYSIS OF PLANT-TYPE VARIABLES
RELATED TO YIELD OF DRY BEANS
(PHASEOLUS VULGARIS L.)

By

Jacques Clarel Denis

Negative correlations between the components of yield, namely number of pods per plant (X), number of seeds per pod (Y) and seed weight (Z), are often seen as major barriers to yield improvement in beans. A question is raised as to whether these traits are part of a larger set of variables which determine collectively the pattern of production of the bean plant. In trying to answer this question, an image factor analysis was carried out on a set of twenty-two variables involved in fixing the yield potential of the plant.

The material consisted of sixteen varieties, of which eight are determinate, eight indeterminate or six are light-seeded, five have medium seed weight and five are heavy-seeded lines. They were grown at two locations in Michigan during the summer of 1970. A randomized block design with four replications and single-row plots was used at both locations. But, in order to accentuate the differences between the two environments, a standard number of sixty seeds per plot was sown at East Lansing, the first location, whereas a standard weight of fifteen grams was planted at Gratiot, the second location. The rows were always

6 meters long and 50 centimeters apart. Observations were made at maturity and all analyses were performed on the logarithms of the averages over five plants chosen at random in every plot. The twenty-two traits obtained from these observations and included in this study were : total number of nodes per plant, number of nodes per plant with pod-bearing branches, total number of branches per plant, number of racemes on the branches per plant, number of pods on the branches per plant, total number of racemes per plant, total number of pods per plant, average number of seeds per pod, average seed weight, average number of pods per productive node, average plant weight, average pod weight, average pod breadth, average pod thickness, average pod length, number of bassal or short internodes, number of upper or long internodes, average short-internode length, average long-internode length, average hypocotyl diameter, average short-internode diameter, average long-internode diameter, numbered in the same order they are listed here. A full model of factor analysis was assumed and squared multiple correlations were used as communalities. The correlation matrix with these communalities as elements of the leading diagonal was then adjusted and became the image covariance matrix, the input to image analysis. The analyses were performed on total variation, and only the factors with eigenroot greater than unity were extracted. The results obtained with the complete-set of data were checked through analyses of sub-sets of data, namely analyses by location, by growth habit, by seed-weight and double-checked through the use of multiple discriminant analysis.

Three major factors or patterns of production were found and together they accounted for 83% of the total variation before rotation and about 77% after rotation. The first two were about evenly important and each extracted after rotation 31% of the variance. The first factor or pattern of production was identified as pod-weight, and characterized a plant-type that will reach a high yield through a high seed weight. This plant has long and sturdy basal internodes, no or very few long internodes and it produces very large pods. The second pattern of production identified a plant-type capable of reaching a high yield through a high number of pods per plant. Like the pod-weight type, the pod-number type has no or very few long internodes but it has instead many moderately sturdy short internodes. Too many nodes will provoke a reduction in the yield of either type. It has been noted that although these two types can be inproved to yield more, far better yield can be obtained with intermediate plant-types resulted from inter-breeding between them. New varieties can thus be produced that will approximate the ideals of the plant breeder for maximum yield on a per-plant basis or an area basis, depending on the number of branches judged acceptable by assuming that the greater their number, stronger a competitor the plant might be. The third factor, a growth factor, characterized by an excessive vegetative growth coupled with a very low production, was seen as a limit to this inter-breeding.

The factors leading to these conclusions are statistically uncorre- but they are not believed to be so from a biological standpoint. So, it

would be very interesting to submit these factors to an oblique rotation and determine their natural association. This can be seen as an important next step, in view of its implication on the possibility of combining maximally the characteristics of the two major plant-types encountered here.

FACTOR ANALYSIS OF PLANT-TYPE VARIABLES RELATED TO YIELD OF DRY BEANS (*PHASEOLUS VULGARIS L.*)

By

Jacques Clarel Denis

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Crop and Soil Sciences

1971

This thesis is dedicated to those whose labor or thoughts have contributed to the advancement of Science in general, and Agronomy in particular, but whose names, at some point in history, for one reason or another, have ceased to be remembered.

ACKNOWLEDGEMENTS

The author expresses his gratitude and appreciation to Dr. M. W. Adams, the chairman of his advisory committee, for his guidance and support at every stage of the realization of this study.

His sincere thanks go to Drs. C. E. Cress, J. W. Hanover and W. L. Myers, as members of his guidance committee, for their much appreciated help.

A special thank to Mr. Brian L. Doyle, for the patience and care he put in helping me obtain the data.

To my wife, Wilberte, my children, Nancy and Charles, for their love and understanding, I want to say : thank you and may God bless you.

Finally, I wish to acknowledge the financial support I received from the Department of Crop and Soil Sciences, without which my program of study would not probably even be started.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF TABLES AND FIGURES (Cont.)

INTRODUCTION

The yield plateau of the present varieties of dry beans, *Phaseolus vulgaris* L., is low compared to that of many other field crops. Although several reasons for low yields in beans can be adduced, I shall mention two problems that are seen as major barriers to yield improvement: one, the physical limitation of the present bean plant-types; and two, the faulty synchionization that often prevails between the demands of the plant for environmental resources and the availability of these resources to the plant over time.

Bean varieties are either bush or vine. Both types can produce a relatively high number of flowers but retain a much reduced number after fertilization. Of these retained and transformed into pods, still fewer will reach the harvest stage. This explains in part the close association generally observed between the number of pods per plant (X) and the total seed yield (W). Seldom can this association be exploited in a plant yield improvement program however, since along with it there exist also negative correlations between the yield components, namely the number of pods per plant (X), the number of seeds per pod (Y) and the average seed weight (Z); here yield is taken as the product of these three components. Some of the questions that can be raised at this point are: Is it the waste of nutrients occasioned in part by the excessive production of flowers which is responsible for the negative correlations between yield components, thus implying some kind of limit on food

availability together with a marked competition between the components for that food? Or, is it instead a differential through time in the amount and quality of environmental resources available and/or in the photosynthetic capacity of the plant which is the main reason for both the high correlation between X and W and the negative correlations between the components? These two questions are relevant particularly to the second problem mentioned above. But, one can also ask: Are these components the only variables involved or are they only part of a much larger set of traits which together determine the pattern and level of production of the bean plant. To answer this question is to determine whether these patterns exist or not and if so, how they are associated with the different bean plant types. This is the main objective of the present work and it is seen as a first step in the understanding of the problem of physical limitations of the present bean plant types, referred to above. Factor analysis is the statistical tool that will be used to this end.

REVIEW OF LITERATURE

Factor analysis is a statistical technique which facilitates the interpretation of the interrelationships among many variables (Harman, 1968). It consists in the reduction of a large number of observed correlated variables to a much smaller number of hypothetical uncorrelated variables called factors. Scientifically, these factors or constructs are the basic influences or dimensions in the data which can be used later instead of the more numerous original variables since they explain most of the correlations among them (Cattell, 1965).

There exist many methods of factor analysis. They will differ depending on the assumptions made or the model chosen and the specific goal being pursued. However, on the basis of their order of execution, they can be divided between direct solutions methods and derived solutions methods (Harman, 1967). The direct solutions methods would be those used to transform a correlation matrix into a factor matrix, i.e., to find the common-factor space of that correlation matrix. The derived solutions methods concern a reorientation or rotation of the reference axes of that common-factor space, so as to facilitate the interpretation of the results, hence to approach a "simple structure" representation. In other words, the direct or unrotated solution defines the pattern of variance in the sample data, whereas the derived or rotated solution reveals the clusters of intercorrelations among the variables (Rummel, 1970). In the present work, a version of principal-factor analysis, image analysis, and varimax orthogonal rotation have been used as direct

solutions and derived solutions methods, respectively.

Principal-factor analysis as performed today on electronic computers has been developed by Kelley (1935) from the method of principal components of Hotelling presented in 1933. The first computer application was made by Wrigley and Neuhaus in 1952. The original method itself has its foundation in the "method of principal axes" of Karl Pearson (1901), where he showed how a large set of data can be reduced empirically so that a maximum portion of the variance present in the data can be accounted for. Hotelling, in turn, made of this technique the principal components method. However, the father of factor analysis is considered to be Charles Spearman who dedicated forty years of his life to its development (Harman, 1967, 1968). Principal components analysis, as mentioned previously, is a method by which many correlated variables which together can possess some degree of redundancy are reduced to a few uncorrelated variables (the principal components), thus eliminating that redundancy (Bailey, 1956). There will be as many principal components or factors as there are variables in the analysis, although only a few are needed to account for a large part (90% or more) of the sample variance. All of these components should be included, however, in order to reproduce the correlations among the variables. According to Maxwell and Rao, cited individually by Murty and Arunachalam (1967), no assumption is made concerning the distribution of the variables or their randomness. In principal-factor analysis, however, it is assumed that the observed variables follow a multi-variate normal distribution and that the newly established factors (fewer in number than the variables) are linear and additive.

The principal components analysis model is:

$$Z_j = a_{j1} F_1 + a_{j2} F_2 + \cdots + a_{jn} F_n \qquad (j = 1, 2, \cdots, n)$$

where the observed variable Z_j, is expressed as a linear regression on _n_ factors (F's). The factors are considered as random variables and they can be defined by a normal probability density function (Harman, 1967). Harman referred to this model as a "maximum variance model". By contrast, the classical factor analysis model is constructed so as to reproduce most of the correlations among the variables. It can be written as follows:

$$Z_j = a_{j1} F_1 + a_{j2} F_2 + \cdots + a_{jm} F_m + d_j U_j \quad (j = 1, 2, \cdots, m),$$

where each observed variable is described as a linear regression on _m_ factors only (_m_ smaller than _n_) plus a unique factor (U_j). The common factors result from the correlations among the variables whereas the unique factor represents the remaining variance (including error) of a particular variable not accounted for in the set of correlations. A good distinction between the implications of the two models is explicit in this characterization by Cattell, who called the principal component model a closed model and the principal-factor one an open model. In practice however, the method of "principal factors" is known as an application of principal components analysis to a reduced correlation matrix, that is a matrix which has communalities in its main diagonal instead of unities (Cattell, 1965; Harman, 1967). The rest of the correlation matrix is further adjusted in image analysis (Veldman, 1967, Rummel, 1970).

The communalities can be estimated in many ways. Rummel discussed at some length three categories of them. He reported also the squared multiple correlation as being the most commonly used communality estimate (Rummel, 1970). As noted by Harman, Thomson (1934) was the first to use Hotelling's principal components method with the full model

of classical factor analysis (Harman, 1967)

Factor analysis has a long list of very diverse fields of application which includes international relations, sociology, economics, communications, taxonomy, biology, geology, meteorology and particularly psychology and the behavioral sciences (Rummel, 1970, Harman, 1967, 1968). However, most of these applications are as recent as the availability of compouter programs. In the last four or five years, a few works related to plant breeding have been published. For example, in 1967, Murty and Arunachalam, using factor analysis found that the pattern of divergence in the genus Sorghum differs under natural and artificial selection (Murty and Arunachalam, 1967). Later, joined by Jain, they established that, this was the case for both cross and self-pollinated crops. They also determined that selection, be it natural or artificial, has more influence on the factor loadings than the breeding system. Of greater importance was their finding that for both outbreeders and inbreeders, under natural or artificial selection, the variables associated with productivity were the ones with highest loadings on the first factor (Murty, Arunachalam and Jain, 1970). Morishima, Oka and Chang, working with samples from an F_3 population of a cross between a tall and a short strain of rice, applied principal component analysis to a matrix of average genetic correlations between seven morphological traits. After rotation to "simple structure", they identified two bipolar axes of variation in rice plant type, namely the panicle-number versus panicle-length and the internode-length versus internode-number axes. The panicle-number and the internode-length types had higher yielding potential (Morishima, Oka and Chang, 1969). Using the same material, Hashigushi and Morishima performed a principal components analysis on

the phenotypic correlations. They used their own method to determine from the results what they called "the genetic vectors". Interestingly, and significantly, these were comparable to the components extracted from the genetic correlations (Hashigushi and Morishima, 1969). Hegman and DeFries found different results with phenotypically standardized phenotypic, genetic and environmental covariances than with the unstandardized covariances (Hegman and DeFries, 1970). This shows clearly that scaling can have an effect on results from factor analysis. The differences in size of loadings in the principal components extracted from genetic and environmental correlations, in the work by Murty and Arunachalam, seem to indicate the importance of normality as a requirement in factor analysis (ibid.). Consequently, these authors express the view that the environmental correlations are more appropriate for factor analysis.

MATERIALS AND METHODS

Twenty-four homozygous lines of dry edible beans were planted at two locations in Michigan during the summer of 1970. Single-row plots in a randomized block design with four replications were used. The rows were 6 meters long and 50 centimeters apart. In order to accentuate the differences between the two environments, a standard number of seeds (sixty per plot) was sown at East Lansing, the first location, and a standard weight of seeds (fifteen grams per plot) was planted at Gratiot, the second location. I believe that the light-seeded lines are usually favored by the first, more popular type of planting, so I provoked deliberately the opposite situation in the second type of planting. Observations were to be recorded at maturity, because it is the proper time to have a measurement of some of the traits in which I was interested. As an indication of maturity, it was taken that most of the plants in the plot would have to have almost all the pods fully developed and, in addition, 25 to 50% of them should be at the pale green or yellowish color stage. Because of frost damage to different varieties at the two locations, complete sets of data have been obtained on a smaller number of entries than anticipated. Consequently, the present work includes sixteen varieties for which all data have been recorded at both locations. These varieties are listed below. Eight are bush (B) or determinate and eight are vine (V) or indeterminate, or six are

light-seeded (L), five have medium seed weight (M) and five are heavy-seeded (H) lines. Incidentally, the light-seeded varieties correspond to navy beans, the medium seed weight to great northern beans and the heavy-seeded lines to kidney beans.

The sixteen varieties were:

1.	02-	Great Northern 1-27	(M, V)
2.	03-	Algarrobo	(H, B)
3.	04-	Sanilac	(L, B)
4.	05-	Charlevois	(H, B)
5.	08-	Criolla	(L, B)
6.	09-	Dominican	(L, V)
7.	10-	Mich. Cranberry	(H, V)
8.	12-	Red Mexican	(M, V)
9.	13-	Michelite-62	(L, V)
10.	14-	Perry Marrow	(M, V)
11.	15-	Seafarer	(L, B)
12.	16-	Pinto 114	(M, V)
13.	17-	Manitou (RKB0255)	(H, B)
14.	18-	Rinson Oscuro	(M, B)
15.	21-	Saginaw	(L, V)
16.	23-	Estrada Rosada	(H, B)

At maturity, five plants per line were chosen at random and observations were made on each plant on a node per node basis. However, all analyses were performed on logarithms of the averages over the 5 plants because of the disparity in the variances of the different variables.

Twenty-two traits were finally retained after elimination of complex traits like total seed yield and number of seeds per plant which would probably make the results less interpretable. Also, some traits were not used in order to prevent singularity or too low communality in the final set of traits. The twenty-two traits or variables together determine yield, the coefficient of determination exceeded 98%. These variables were:

1. total number of nodes per plant
2. number of nodes with pods on branches per plant
3. total number of branches per plant
4. number of racemes on the branches per plant
5. number of pods on the branches per plant
6. total number of racemes per plant
7. total number of pods per plant (X)
8. average number of seeds per pod (Y)
9. average seed weight (Z)
10. average number of pods per productive node
11. plant fresh weight
12. average pod fresh weight
13. average pod breadth
14. average pod thickness
15. average pod length
16. number of short-internodes*

---

*Those immediately above the hypocotyl (5 or 6 on the average), which are followed in turn by the long-internodes, more variable in number.

17. number of long-internodes
18. average short-internode length
19. average long-internode length
20. average hypocotyl diameter
21. average short-internode diameter
22. average long-internode diameter

Principal-factor analysis comprises two main parts, namely the calculation of the correlation matrix (R) and the extraction of the eigenvalues and eigenvectors of that correlation matrix. In factor analysis, the use of the correlation matrix as calculated implies that all the correlations result from common factors or common variance. Consequently, neither specific factor of a variable nor error is accounted for. Such an assumption has been discarded in this work. So, the image covariance matrix was used instead of the correlation matrix, thus following the full model of classical factor analysis. The image covariance matrix is a transformed correlation matrix in which the unities in the main diagonal have been replaced by the communalities of the variables, here their squared multiple correlations, and then the rest of the matrix is adjusted accordingly. Then the Jacobi method was applied to the image covariance matrix to extract its eigenroots and eigenvectors. The eigenroots or eigenvalues are the variances or portions of the total variation extracted by the corresponding eigenvectors or principal components. Not all the components can be considered as important, since the first ones are always associated with a greater part of the total variation. So, an eigenroot of at least one was chosen as requirement for an eigenvector to be extracted. After

extraction, the factors were submitted to varimax orthogonal rotation. Varimax as opposed to quartimax rotation simplifies the factor instead of the variables, and this is done so that the orthogonality between any two factor-vectors is preserved.

All analyses were made using the total variation in the sample and assuming the full model of factor analysis. The discussions of the results will concern only the common part of that variation. Beside the complete set of data, analyses were also performed on sub-sets of the data, namely, each location separately, each growth type separately, and finally each seed-weight group separately. The factors extracted for groups within each classification were then compared. The varimax factor loadings were used as bases for these comparisons. Two types of results were obtained here, the correlations between the factor-variables in the two sets of factors being compared, and the correlations between all pairs of variables in the two sets (same variables in both). The diagonal elements of this last matrix are correlations of a given variable in a particular sub-set with the same variable in the other sub-set. Prior to the calculation of these correlations the second structure or set of factors is generally transformed to attain maximum overlap with the first. Finally the results of these comparisons are double-checked through the use of multiple discriminant analysis. Since this analysis can indicate which variables in a set account for differences between groups of subjects, when these differences exist, it should be possible, through it, to verify the conclusions arrived at after finding and comparing the rotated factors.

The computer programs used in this work are part of a package offered by D. J. Veldman in his book "Fortran Programming for the Behavioral Sciences" (Veldman, 1967).

RESULTS

Factor Analysis of the Complete Set of Data

The results from the complete set of data indicate that 92.70% of the total variation is common, that is, the common factors can account for that much of that variation, and the remaining 7.30% should be attributed to unique factors and errors. The trace was 20.4, 83.16% of which was extracted by only 3 roots. The first two roots accounted for more than 73% before rotation and more than 62% after rotation (Table 1, page 15). So, it can safely be said that there are mainly two factors at work in the present set of data. The varimax rotation results show the third factor increasing in importance (from about 10% before rotation to about 15% after), whereas the first two become about evenly important (31% each).

The variables with the highest loadings in the first factor are: average pod thickness (14)*, average seed weight (9), average pod fresh weight (12), average pod breadth (13); next in importance are: average pod length (15), average short-internode length (18), all three diameter-variables (20-22) and plant fresh weight (11) (Table 1 and Figure 1 on page 15). Number of seeds per pod (8) has a moderately high and negative loading in this factor. The variables with the lowest loadings are: for the positive, number of nodes with pods on the branches (2), number of

---

*The numbers in parentheses identify the variables as listed on pages 10-11.

Table 1. Factor Analysis Results, Complete Set of Data

		Unrotated Factors				Rotated Factors		
Traits	% C*	1	2	3	%C*	1	2	3
1	97.61	-.6237	.3475	.6117	88.39	-.3250	.2509	.8458
2	97.24	.1393	.8606	.1122	77.25	.0364	.8739	.0867
3	96.26	.2890	.8062	.0306	73.45	.1394	.8437	-.0565
4	99.64	.1676	.9337	.1128	91.27	.0499	.9508	.0788
5	99.49	.1598	.9539	-.0798	94.19	-.0473	.9653	-.0883
6	99.44	-.1459	.9241	.1992	91.50	-.1844	.8906	.2963
7	99.42	-.1965	.9184	-.0736	88.75	-.3515	.8708	.0756
8	86.24	-.3906	.4619	-.3270	47.29	-.5641	.3830	-.0891
9	96.68	.9183	-.1573	.2317	92.17	.9355	.0036	-.2158
10	91.22	.3075	.6357	-.4277	68.16	-.0249	.6691	-.4830
11	94.51	.6510	.5728	.1286	76.85	.5385	.6763	-.1455
12	98.68	.9102	-.1082	.2403	89.79	.9245	.0508	-.2016
13	92.38	.8268	-.2223	.3254	83.90	.9080	-.0738	-.0948
14	96.39	.8527	-.1443	.3651	88.12	.9363	.0082	-.066[illegible]
15	91.60	.8407	-.0659	-.0157	71.13	.7408	.0755	-.3961
16	76.56	-.3571	.5201	.0074	39.81	-.3924	.4530	.1971
17	97.59	-.6746	.2596	.6119	89.69	-.3556	.1557	.8638
18	80.80	.6273	-.3283	.2595	56.86	.7200	-.2131	-.0699
19	67.71	-.3047	-.0189	.5800	42.96	-.0014	-.0575	.6529
20	95.73	.8649	.3333	-.0586	86.26	.6792	.4722	-.4223
21	96.54	.7904	.4386	.0122	81.72	.6292	.5649	-.3197
22	87.56	.8219	-.0134	-.2952	76.28	.5893	.1182	-.6337
E.Roots		8.1827	6.7969	1.9783				
% Trace		40.13	33.33	9.70	% Var.	31.28	31.03	14.78
Cum.% Trace		40.13	73.46	83.16	Cum % Var.	31.28	62.31	77.09
Total Trace = 20.3931								

92.70% of total variation is common.

* Communality

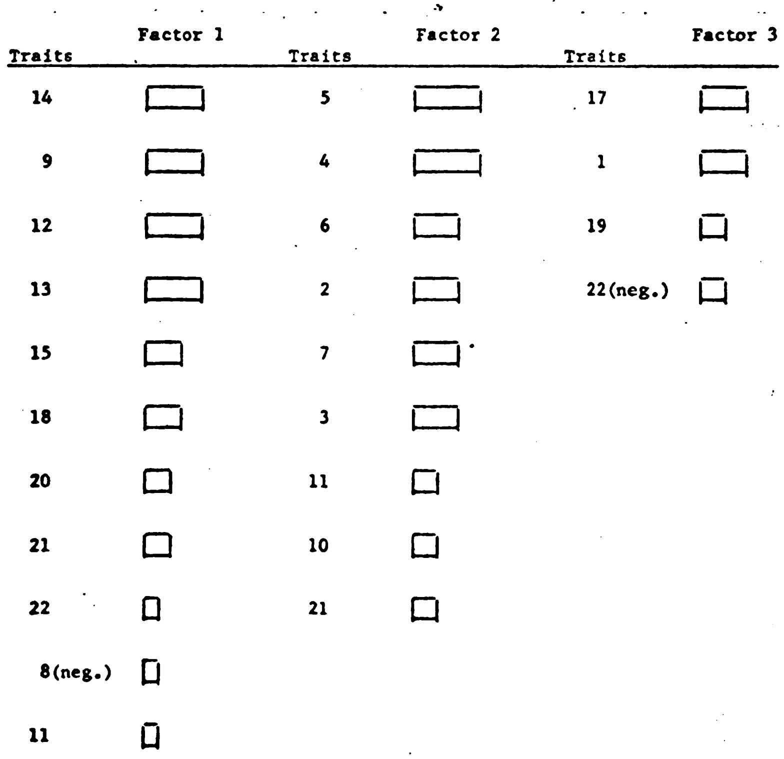

Figure 1. Comparison of size of loading for the variables included in the three main rotated factors.

racemes on the branches (4), and total number of branches (3); and for the negative, number of pods on the branches (5), average number of pods per productive node (10) and average long-internode length (19). Essentially this is a weight factor.

The second factor is highly associated with number-variables like number of nodes with pods on the branches (2), total number of branches (3), total number of racemes (6) and number of pods per plant (7), particularly with number of racemes on the branches (4) and number of pods on the branches (5), but also with number of pods per productive node (10) and average short-internode diameter (21) (Table 1 and Figure 1 on page 15). In general, the uppermost three or four short or basal internodes support most of the productive branches of the bean plant. Seed weight (9), pod weight (12), pod thickness (14) and pod length (15) have the lowest positive loadings whereas pod breadth (13) and long-internode length have the lowest negative loadings in this factor. Consequently, this factor indentifies itself with numbers, particularly number of branches, but not with weight nor with large pods.

The third and least important factor is characterized by loadings which reflect a numerous-thin-long internode relationship, that is, total number of nodes on the plant (1), number of long-internodes (17) and average long-internode length (19) have the highest positive loadings in this factor whereas the diameter-variables (20-22) along with average number of pods per productive node (10) and average pod length (15) have the highest negative loadings (Table 1, Figure 1 on page 15). Number of nodes with pod-bearing branches (2), number of racemes on the branches (4) and number of pods per plant (7) have the lowest positive loadings. Total

Table 2. Factor Analysis Results, East Lansing Set

		Unrotated Factors					Rotated Factors			
Traits	% C*	1	2	3	4	% C*	1	2	3	4
1	97.39	.5829	-.2427	.6629	-.2341	89.29	.1936	.1334	.9151	.0144
2	97.66	.5788	.6758	.0551	.1095	80.67	.1251	.8778	.0448	.1360
3	95.95	.4990	.6385	.0225	.2026	76.46	.1107	.8397	-.0494	.2117
4	99.77	.6386	.7000	.1141	-.1803	94.33	.0223	.9442	.2089	-.0870
5	99.59	.6845	.7021	-.0040	-.0481	96.39	.1485	.9658	.0877	-.0374
6	99.73	.8129	.4811	.2218	-.0822	94.83	.2456	.8554	.3954	-.0031
7	99.74	.8620	.4362	.0287	.0796	94.05	.4343	.8384	.2193	.0284
8	91.10	.6354	-.0773	-.3160	.5184	77.83	.8043	.2670	-.1484	.1950
9	95.92	-.7878	.4763	.1621	-.2342	92.86	-.9289	-.0338	-.2501	.0461
10	95.36	.5021	.6624	-.2395	-.1261	76.42	.0923	.8351	-.1351	-.2002
11	93.41	.0378	.7870	.1314	.3516	76.17	-.2444	.6574	-.2308	.4652
12	98.76	-.8175	.4147	.1828	.1549	89.77	-.7780	-.1191	-.3621	.3837
13	94.59	-.8321	.2964	.2793	.0452	86.03	-.8120	-.2204	-.2191	.3231
14	97.35	-.7655	.4142	.2667	-.1676	85.68	-.8967	-.0760	-.1629	.1427
15	92.04	-.7060	.4067	-.0804	.2066	71.30	-.5790	-.0655	-.5411	.2840
16	84.90	.7913	-.0091	-.0582	-.1007	63.97	.5493	.4380	.3135	-.2186
17	96.93	.5502	-.2615	.6748	-.1922	86.35	.1909	.0976	.9023	.0577
18	85.08	-.6883	.2439	.3969	.1692	71.93	-.6801	-.1896	-.0947	.4602
19	75.75	-.0983	.2380	.5455	.4970	61.10	-.1841	.1193	.1523	.7346
20	96.28	-.2282	.8564	-.1414	-.1634	83.21	-.5665	.5904	-.3983	-.0629
21	97.40	-.2045	.8495	-.1254	-.2615	84.76	-.5906	.6023	-.3401	-.1428
22	88.61	-.6009	.5502	-.3526	-.1600	81.37	-.6134	.1289	-.6297	-.1556
E-Roots		8.7944	6.2245	1.9697	1.1592					
% Trace		42.42	30.02	9.50	5.59	% Var.	28.18	31.85	15.31	7.15
Cum. % Tr.		42.42	72.44	81.94	87.53	Cum. % Var.	28.18	60.03	75.34	82.49

Total Trace = 20.7331

94.24% of total variation is common.

* Communality

Table 3. Factor Analysis Results, Gratiot Set

		Unrotated Factors				Rotated Factors		
Traits	% C*	1	2	3	% C*	1	2	3
1	98.86	-.4887	.7250	.4063	92.95	-.3286	.3457	.8379
2	98.79	.4626	.7598	-.0133	79.15	.1391	.8756	.0745
3	98.42	.6613	.5699	-.0485	76.45	.3209	.8041	-.1220
4	99.76	.5690	.7705	.0237	91.80	.2390	.9260	.0579
5	99.65	.6164	.7143	-.2501	95.28	.1289	.9461	-.2025
6	99.56	.2210	.9076	.1479	89.45	.0101	.8688	.3737
7	99.37	.2052	.8667	-.2417	85.18	-.2204	.8944	.0576
8	91.77	-.1499	.5247	-.4556	50.53	-.5244	.4732	-.0804
9	99.45	.8866	-.2781	.3175	96.42	.9373	.0969	-.2762
10	95.14	.6526	.0058	-.6041	79.08	.1394	.4032	-.7803
11	98.24	.8783	.3200	.0637	87.79	.6198	.6623	-.2347
12	99.11	.9135	-.1613	.2812	93.95	.9048	.2177	-.2710
13	95.45	.8034	-.2399	.3749	84.35	.8975	.0831	-.1766
14	97.68	.8239	-.1879	.4388	90.67	.9367	.1266	-.1152
15	96.86	.8790	-.1746	.0629	80.71	.7535	.2291	-.4322
16	84.51	-.0957	.5511	.2996	40.26	-.0453	.3878	.5002
17	98.84	-.6014	.6539	.3801	93.38	-.4107	.2376	.8418
18	87.31	.5576	-.4545	.3038	60.98	.7264	-.2025	-.2027
19	93.22	-.6039	.3272	.5004	72.22	-.2534	-.0713	.8080
20	98.44	.9540	-.0339	.0366	91.28	.7576	.3888	-.4331
21	98.54	.9201	.2024	.0185	88.78	.6567	.5858	-.3366
22	95.34	.8110	-.2437	-.3129	81.50	.4989	.2041	.7241
E. Roots		10.0829	5.8575	2.0810				
% Trace		47.46	27.57	9.80	% Var.	31.80	29.94	20.18
Cum % Trace		47.46	75.03	84.83	Cum. % Var.	31.80	61.74	81.92

Total Trace = 21.2432

96.56% of total variation is common.

* Communality

number of branches (3), number of pods on the branches (5), number of seeds per pod (8), average pod breadth and thickness (13, 14) and finally average short-internode length (18) have the lowest negative loadings in this factor. Therefore, this factor expresses a growth type where at the extreme an exagerate vegetative growth causes a very low production.

Factor Analysis of Individual Locations

The analyses by location show that 94.24% of the total variation is common at East Lansing, that is, the common factors accounted for that much variation (Table 2, page 17) compared to 96.56% at Gratiot (Table 3, page 17), indicating the presence of a slightly greater error at the first location. The results from these analyses were identical to those obtained with the complete set of data, with only one insignificant difference that the signs of the loadings in the first factor at East Lansing were inverted. That should not change the interpretation of the results, however. Another finding was the extraction of a fourth factor at East Lansing, confirming the fact already mentioned that there was more variance at that location. Moreover, the comparison between the four factors of East Lansing with the three of Gratiot gives the following correlations between the corresponding first three factors: -.93, .99 and .93 with no large values encountered for the fourth factor at East Lansing. But, the size of the correlations of that factor with the first and third of Gratiot (.33 and .37 respectively) may explain in part the relatively smaller correlations found for the first and third sets of factors (Table 3 a-b, page 19).

Table 3a. Factor Relationships Between East Lansing and Gratiot Analyses (Correlations)

East Lansing	Gratiot 1	Gratiot 2	Gratiot 3
1	-.9315	.1307	.0329
2	.1147	.9894	-.0656
3	-.0915	.0498	.9264
4	.3329	.0390	.3692

Table 3b. Relationships Between Identical Pairs Of Variables From East Lansing and Gratiot Analyses

Variables	Correlations
1	.9476
2	.9858
3	.9268
4	.9559
5	.9571
6	.9739
7	.9874
8	.8213
9	.9451
10	.6710
11	.8962
12	.8686
13	.8931
14	.9328
15	.8584
16	.6379
17	.9534
18	.9070
19	.3121
20	.9721
21	.9570
22	.9605

Factor Analysis of Determinate and Indeterminate Varieties as Sub-Sets

The results of the analyses by growth type show that there were four factors with roots greater than unity for the determinate lines and five factors for the indeterminate lines. There was greater error variance among the indeterminate lines than the determinate varieties as indicated by the amount of the total variation which is common in the two situations (92.01% compared to 95.63%). Here again, two major factors were recovered, but with some important changes in each case.

For the determinate lines, there was less emphasis put on average short internode length (18) but more on number of nodes with pods on the branches (2), on number of branches per plant (3) and also on number of racemes on the branches (4) (Table 4, page 21). The loadings for the three diameter traits (20, 21, 22) were markedly higher. The second factor was almost identical to that of the complete data set, with a little less importance given to the number of nodes with pods on branches (2) and number of branches per plant (3). These two factors account for more than 76% of the total variation before rotation and for about 64% after. The third factor was also recovered, but it can not be said that it is more important than the fourth factor of the bush-type structure, a factor which seems to contrast short internode length (18) with the number of short internodes (16) and also the number of seeds per pod (8).

For the indeterminate lines, the first factor was identical to the second factor of the complete set, whereas the second resembled the first factor of the complete set of data, but with two major differences. The average long internode length (19) has gained in importance and the diameter variables (20-22) were transposed from the second to the third

Table 4. Factor Analysis Results, Determinate Lines

		Unrotated Factors					Rotated Factors			
Traits	% C*	1	2	3	4	% C*	1	2	3	4
1	98.46	-.6377	.5966	.3880	.1131	92.59	-.5310	.1235	.6672	-.4285
2	98.77	.5542	.7116	.3518	-.0412	93.90	.4388	.6992	.2354	-.4495
3	96.44	.6259	.6137	.3004	.0166	85.89	.5000	.6819	.1883	-.3293
4	99.56	.7489	.5847	-.0515	.1255	92.12	.4472	.8449	-.0481	-.0716
5	99.28	.6805	.6654	-.2089	.1578	97.44	.2943	.9357	-.1066	-.0273
6	99.15	.3063	.8690	-.1935	.1760	91.74	-.0718	.9393	.0687	-.1586
7	99.36	.1248	.8930	-.3367	.1464	94.80	-.2958	.9147	.0047	-.1540
8	91.90	-.2809	.5979	-.1887	-.5691	79.60	-.4588	.2632	-.2496	-.6737
9	98.73	.8939	-.3502	.1468	.1270	95.94	.9131	.0937	-.1557	.3045
10	93.63	.6369	.2014	-.6406	.1700	88.54	.1862	.6725	-.5078	.3751
11	96.62	.8139	.3825	-.0201	-.0810	81.58	.5810	.6442	-.2086	-.1408
12	99.14	.9223	-.2128	.1326	-.1493	93.58	.9065	.1462	-.3023	.0363
13	94.17	.8489	-.2759	.1671	-.0388	82.61	.8718	.0833	-.2066	.1283
14	98.14	.9010	-.2763	.0418	.0235	89.04	.8562	.1593	-.2774	.2344
15	96.20	.7822	-.2528	.2319	-.1839	76.33	.8453	.0147	-.2185	-.0288
16	89.42	-.1186	.5608	.3722	-.5113	72.84	-.0591	.1535	.1387	-.8259
17	97.77	-.6787	.4176	.3488	.3636	88.89	-.5470	.0407	.7540	-.1398
18	89.57	.4655	-.5677	.0333	.3109	63.67	.5486	-.1835	-.0590	.5465
19	77.44	-.4659	.1839	.3521	.5120	63.70	.3088	-.0172	.7278	.1076
20	97.53	.9538	.0600	.1120	.0702	93.09	.8397	.4459	-.1382	.0888
21	98.05	.9242	.1048	.1995	.0060	90.49	.8452	.4261	-.0925	-.0215
22	94.58	.8880	-.2276	.0469	-.0046	84.25	.8357	.1839	-.2751	.1862
E. Roots		10.6408	5.4162	1.5705	1.2988					
% Trace		50.58	25.74	7.46	6.17	% Var.	38.03	26.14	11.02	10.84
Cum % Tr.		50.58	76.32	83.78	89.96	Cum. % Var.	38.03	64.17	75.19	86.03

Total Trace = 21.0392

95.63% of total variation is common.

* Communality

Table 5. Factor Analysis Results, Indeterminate Lines

		Unrotated Factors						Rotated Factors				
Traits	% C*	1	2	3	4	5	% C*	1	2	3	4	5
1	82	.4444	.1540	.5118	-.0241	.0219	48	.2031	-.1796	.6225	-.0914	.1219
2	98	.7608	.2408	-.4287	-.2300	.1171	89	.9087	-.0983	-.1076	-.0575	.1921
3	98	.7512	.2992	-.4011	-.2446	.1137	89	.9167	-.0478	-.0660	-.0696	.1888
4	100	.8735	.3463	-.2523	-.1000	-.1445	98	.9636	-.1398	.1604	.0499	-.0364
5	100	.9300	.2624	-.2064	-.0319	-.0463	98	.9335	-.2449	.1833	.0960	.0733
6	100	.8860	.3782	-.0540	-.0387	-.2081	98	.8976	-.1641	.3624	.0710	-.0827
7	100	.9517	.2114	.0145	.0683	-.0568	96	.8197	-.3455	.3759	.1381	.0832
8	93	.5894	-.1770	.0658	.1032	.7094	90	.2858	-.4120	.0975	.0872	.7927
9	96	-.6406	.6726	-.0011	.0185	-.1501	89	-.2196	.8730	.0864	.0862	-.2457
10	95	.8483	.1533	-.0353	.0439	-.0450	75	.7393	-.3269	.2774	.1121	.0763
11	93	.3503	.8359	.0196	-.0802	.1769	86	.6097	.5252	.3999	.0568	.2210
12	99	-.6053	.7391	-.0200	-.0998	.1780	95	-.1608	.9587	.0559	-.0122	.0803
13	96	-.6969	.6297	-.0980	-.1719	.1258	94	-.2269	.9336	-.0823	-.0851	.0099
14	98	-.6404	.6956	-.0884	-.1363	-.1929	96	-.1404	.9213	.0043	-.0406	-.2957
15	92	-.3129	.6677	.3021	-.1291	.4627	87	-.1110	.7287	.3585	-.1072	.4272
16	86	.6300	-.1562	.3666	-.1639	-.0976	59	.3192	-.5030	.4281	-.2311	.0213
17	78	.1601	.1172	-.1738	.7736	-.0130	67	.0971	-.0660	.0370	.8082	.0048
18	82	-.5157	.6355	-.1651	.0150	-.1942	73	-.0644	.7975	-.0303	.1189	-.2825
19	73	-.2993	.4626	-.3857	.4402	.1281	66	.0178	.5490	-.2049	.5618	.0564
20	90	.4020	.6649	.3086	.1979	-.1923	78	.4372	.2424	.6750	.2393	-.1112
21	94	.5611	.6091	.3483	.2264	-.0196	86	.5033	.1194	.7194	.2575	.0859
22	84	.1024	.2946	.6269	.0193	-.0761	50	-.0563	.0839	.6935	-.0702	-.0168
E.Roots		8.8850	5.1820	1.7811	1.1354	1.0615						
% Trace		43.90	25.60	8.80	5.61	5.24	% Var.	30.48	27.41	12.77	5.77	5.60
Cum % Tr.		43.90	69.50	78.30	83.91	89.15	Cum. % Var.	30.48	57.89	70.66	76.43	82.03

Total Trace = 20.2412

92.01% of total variation is common.

* Communality

factor (Table 5, page 21). Consequently, the third factor becomes an association between the number of nodes (1) and the diameter variables (20-22) instead of the number of long internodes (17) as in the complete-set results. It is a significant change *per se*, because this third factor extracts here about the same amount of variation as it does in the complete-set analysis. The fourth and fifth factors each accounted for less than 6% of the total variance. Number and length of long internodes (17, 19) had high loadings in the fourth factor, and number of seeds per pod (8) showed a high loading in the fifth factor. These variables had also low communalities in this analysis (Table 5, page 21). The first two factors extracted a little more than 69% of the total variation prior to rotation and about 58% after rotation.

Comparison Between the Factor-vectors Extracted in Analyses of Determinate and Indeterminate Varieties

The factor-variables from the determinate and the indeterminate lines were compared and the results show, in addition to the transposition between the first and second factors of the two sets, that factor #3 of the determinate group is similar to factor #4 of the indeterminate group and also that factor #4 of the former is the opposite of factor #5 in the indeterminate group (Table 5a, page 23). Consequently, factor #3 of the indeterminate sub-set is unique to that group and can be seen as another expression of the third factor of the complete-set analysis. The following variables in decreasing order of importance account for most of the differences between the determinate-group and indeterminate-group results: number of nodes on the plant (1), average number of pods per productive node, (10) average long internode diameter

Table 5a. Factor Relationships Between Determinate and Indeterminate Lines Analyses (Correlations)

	Bush			
Vine	1	2	3	4
1	.0498	.9650	-.0765	.0025
2	.8647	-.1197	-.0547	.4007
3	.4740	.1616	.0223	-.4834
4	.0087	.0945	.9744	.1908
5	.1579	-.1399	.2029	-.7546

Table 5b. Relationships Between Identical Pairs of Variables From Analyses of Determinate and Indeterminate Lines

Variables	Correlations
1	.1728
2	.6613
3	.6944
4	.8817
5	.8985
6	.9783
7	.9365
8	.7253
9	.8808
10	.4399
11	.9412
12	.8049
13	.8015
14	.8442
15	.9105
16	.6602
17	.7778
18	.9875
19	.4639
20	.7738
21	.7376
22	.4600

Table 5c. Multiple Discriminant Analysis between Determinate and Indeterminate Lines

Variables	Discriminant Axis Loadings
1	.9418
2	.1810
3	.0003
4	.3076
5	.1535
6	.4906
7	.2750
8	.1289
9	-.2645
10	-.1866
11	-.0694
12	-.2511
13	-.2139
14	-.1413
15	-.3661
16	.3882
17	.8727
18	-.2497
19	.3359
20	-.3193
21	-.1638
22	-.5614
% of Total Variation :	100.00
Chi-Square Value :	317.593
Degrees of freedom :	22
Significance Level :	.0000

(22) and average long internode length (19) (Table 5b, page 23).

These results were also confirmed by the use of multiple discriminant analysis. Total number of nodes on the plant (1) was the main differentiating variable between determinate and indeterminate lines, followed by the number of long internodes (17) and also their diameter (22) (Table 5c, page 24). The discriminant function had a highly significant chi-square.

Factor Analysis for the Three Seed-weight Groups

Attempts were also made to test the invariance of the factors when, instead of classifying by growth type, the material was divided into seed-weight sub-sets. The average seed weight before planting was used for this purpose. Three groups were established; the light-seeded, the medium-weight and the heavy-seeded lines. They corresponded approximately to the navy, the great northern and the kidney beans with an average weight of 15-16 gms per hundred dry seeds, 29-30 gms and 40-45 gms, respectively.

Factor analysis for the light-seeded lines reveals that 93.61% of the total variation was common, and that 89.67% was extracted by 5 roots (Table 6, page 26). The first factor accounted for 46.20% of the trace before rotation and 41.6% after rotation. It was by far the most important factor, since the next two factors together extracted only some 31% of the trace. That factor resembles a combination of the first two factors extracted or identified from the complete set of data. The short internodes and hypocotyl diameters of the first factor were combined with the number-variables of the second factor of the complete set to form the first factor of the light-seeded lines. Three remarks should

Table 6. Factor Analysis Results, Light-Seeded Lines

		Unrotated Factors						Rotated Factors				
Traits	% C*	1	2	3	4	5	% C*	1	2	3	4	5
1	98	.7193	-.4807	.1565	.2863	-.2584	92	.7425	-.2362	.2849	.4619	-.1418
2	98	.8437	-.0300	-.4062	.1763	.0639	91	.8473	.1617	-.3523	.2091	-.0326
3	98	.8053	.0584	-.4352	.0967	.1139	86	.8023	.2085	-.4067	.1005	-.0308
4	100	.9614	-.1634	-.0179	-.0063	.0480	95	.9715	-.0128	.0121	.0258	-.0943
5	99	.9585	-.1015	-.0476	-.0923	.1870	97	.9785	.0151	-.0542	-.1172	-.0181
6	99	.9581	-.0718	.1304	-.0417	.0662	95	.9527	.0577	.1426	-.0721	-.0987
7	99	.9323	.0285	.0497	-.1636	.2211	95	.9340	.1043	.0157	-.2526	-.0287
8	95	.3892	.5188	-.6513	.0593	-.0748	85	.2970	.5501	-.6373	.0494	-.2335
9	91	.1406	-.2468	.8573	.1942	-.1258	87	.1451	-.1359	.9037	.1082	.0365
10	93	.5943	.0450	.0848	-.5285	.3671	78	.6176	-.0624	-.0381	-.6242	.0086
11	95	.7473	.5557	.1233	-.0241	.1600	91	.6590	.6107	.0790	-.2977	-.0804
12	97	.2318	.8367	.2408	.3157	-.0495	91	.0819	.9159	.2439	-.0438	-.0826
13	81	-.0125	.6510	-.1494	.5035	.1737	73	-.0828	.7888	-.1344	.1602	.2391
14	93	.2115	.2923	.7949	.2847	.0080	84	.1485	.4054	.8052	-.0358	.0841
15	90	.1643	.7994	.0053	-.0817	-.1870	71	.0008	.7179	-.0271	-.2483	-.3604
16	93	.6887	.1100	-.1000	-.0620	-.4935	74	.5814	.1608	-.0240	.1401	-.5997
17	98	.5692	-.6258	.1652	.3592	-.0296	87	.6546	-.3560	.2770	.4706	.1394
18	86	-.5787	.1110	.5428	-.1551	.1087	68	-.5789	-.0365	.4662	-.3217	.1434
19	73	-.2093	-.1830	.1249	.3788	.6237	63	-.0788	-.0354	.0969	.0798	.7760
20	97	.8828	.0462	.3324	-.1244	-.1075	92	.8292	.1184	.3420	-.1472	-.2806
21	97	.9133	-.1399	.2238	-.0107	-.0656	91	.8892	-.0032	.2605	.0113	-.1778
22	89	.0865	.2992	.6120	-.3491	.0261	59	.0228	.1549	.5220	-.5249	-.1467
E.Roots		9.5144	3.3311	3.1773	1.3594	1.0856						
% Trace		46.20	16.17	15.43	6.60	5.27	% Var.	41.63	14.42	14.15	7.19	6.54
Cum. % Tr		46.20	62.37	77.80	84.40	89.67	Cum. % Var.	41.63	56.05	70.20	77.39	83.93

Total Trace = 20.5951

93.61% of total variation is common.

* Communality

Table 7. Factor Analysis Results, Medium-Weight Lines

		Unrotated Factors					Rotated Factors			
Traits	% C*	1	2	3	4	% C*	1	2	3	4
1	99.10	-.7367	.5662	.2964	.1021	96.16	-.9312	.0133	-.2691	.1478
2	98.98	.6149	.6719	.1374	-.1794	88.06	.1056	.9248	-.1140	.0356
3	98.57	.7203	.5503	.1182	-.1754	86.64	.2452	.8966	-.0367	.0316
4	99.94	.6111	.7291	.1424	-.2424	98.40	.0719	.9799	-.1354	-.0158
5	99.91	.7053	.6779	.1096	-.1181	98.30	.1900	.9595	-.1266	.1007
6	99.88	.3600	.8531	.2814	-.1702	96.54	-.2334	.9379	-.1631	.0689
7	99.82	.5239	.7858	.2142	.0175	93.81	-.0339	.9219	-.1644	.2452
8	93.62	.4252	.4260	.1197	.6924	85.60	.1379	.4114	-.1419	.8047
9	97.80	.6121	-.7136	.0678	-.0413	89.02	.7185	-.0995	.6016	-.0460
10	98.25	.7634	.4643	-.1518	-.0390	82.30	.4596	.7484	-.2063	.0953
11	97.89	.8982	.1093	.3252	.0807	93.09	.4758	.6913	.3831	.2827
12	98.95	.5973	-.7027	.2312	.1656	93.14	.6360	-.1056	.6932	.1877
13	95.06	-.0395	-.7566	.5534	.0053	88.03	.0024	-.3905	.8530	.0152
14	97.59	.2492	-.7481	.3299	-.4106	89.92	.2985	-.1706	.7916	-.3929
15	98.71	.8185	-.4243	-.0865	.2908	94.20	.8584	.0964	.3204	.3053
16	94.56	-.2755	.7301	-.2037	.2927	73.61	-.3835	.2093	-.6736	.3024
17	99.44	-.7423	.4771	.4253	.0048	95.96	-.9712	.0104	-.1066	.0700
18	77.35	.2738	-.5415	.5018	.0189	62.03	.1778	-.0707	.7589	.0884
19	87.04	-.3858	.0529	.7280	.2516	74.49	-.6615	-.0453	.4196	.3594
20	97.98	.9090	-.2867	-.0392	.0597	9182	.8312	.3121	.3251	.1552
21	98.79	.9424	-.1163	-.0442	.1132	91.63	.7906	.4436	.2287	.2052
22	98.21	.8468	-.2579	-.2892	.0438	86.92	.8951	.2330	.1054	.0516
E·Roots		9.0322	7.3545	1.9931	1.1172					
% Trace		42.46	34.57	9.37	5.25	% Var.	31.43	32.22	18.50	6.47
Cum. % Tr.		42.46	77.03	86.40	91.64	Cum % Var.	31.43	63.65	82.15	88.62

Total Trace = 21.2745

96.70% of total variation is common.

* Communality

be made here. Total number of nodes on the plant (1) was included, so was the number of long internodes (17). The pod dimensions (12-15) along with the average long internode diameter (22) vanished. Finally, the number of short and long internodes (16, 17) had positive loadings instead of negative loadings as in the complete-set results and the opposite occurred for the average short internode length (18). So, this is a general factor pointing toward a plant type with sturdy, short internodes producing many branches full of productive units. There can be many nodes on such a plant. The second factor in the light-seeded sub-set was a pod factor characterized by long, wide and filled pods (12, 13, 15) leading to a relatively heavy plant (11). The third factor expressed a contrast between seed size and weight (14, 9) and seed number per pod (8). Such a type would probably have sturdy long-internodes (22) and also long short-internodes (18).

In the analysis of the medium-weight set, there was a relatively smaller error variance, 96.70% of the total variance was common, and 91.64% was extracted by 4 roots. The first two vectors were about equally important as indicated by the amounts of variance extracted by them, 42.46% and 34.57% before rotation and 31.4% and 32.2% after rotation, respectively (Table 7, page 26). The second factor was identical to the second factor of the complete set of data. The first, however, was apparently a completely new factor. It is bipolar and has large positive loadings on the diameter-variables (20-22), the length and the weight of the pod (15, 12) and seed weight (9), and negative loadings particularly on the total number of nodes (1) and on the number and the length of the long internodes (17, 19). Number of pods per productive

node (10) and average plant fresh weight (11) have positive loadings but intermediate in size. The third factor contrasts the number of short internodes (16) with seed weight (9), the pod dimensions (13-15) and the average short internode length (18). The number of short internodes (16) had negative loading. This factor extracted 9.37% of variance before rotation and 18.5% after rotation. It was the most important third factor encountered in this study (Table 7, page 26).

The analysis of the heavy-seeded sub-set shows that 95.64% of the total variation was common and 89.0% was extracted by 5 roots. Only the first three deserve consideration, however (Table 8, page 29). The first factor was essentially the same as the second factor of the complete set analysis and the second was comparable to the third of those factors, but with some important differences. The negative loadings of the diameter-variables (20-22) and number of pods per productive node (10) were significantly higher in this factor. Finally, the third factor is identifiable with a contrast of plant weight (11), pod weight (12), pod breadth (13) and pod thickness (14). There were no high positive loadings on this factor. The three factors-variables of this set accounted for 35.59%, 26.89% and 11.15% of the total variation before rotation and 26.36%, 24.52% and 14.17% after rotation, respectively.

Comparisons Between the Factor-Vectors Extracted from Analyses of the Seed-Weight Groups

The three sets of factors from the three seed-weight groups were compared two by two. Through these comparisons, it can be seen that the second factor of the complete-set factors was found in the results of all three seed-weight groups, although with some modification in the

Table 8. Factor Analysis Results, Heavy-Seeded Lines

		Unrotated Factors						Rotated Factors				
Traits	% C*	1	2	3	4	5	% C*	1	2	3	4	5
1	99	-.1809	.9505	-.1302	.0722	.1396	98	.2478	.9286	.0711	.1438	.1685
2	98	.5940	.6112	.2779	.0329	.1036	82	.6792	.1556	-.0855	.2401	.5047
3	93	.6702	.4152	.4546	-.1989	.0390	87	.6616	-.0839	.0101	.0649	.6482
4	100	.7870	.4801	-.1473	-.0697	-.3118	97	.9638	.0671	-.1614	-.1037	.0583
5	100	.9126	.2500	-.0802	.0142	-.2782	98	.9374	-.1889	-.2491	-.0176	.0507
6	100	.6619	.6499	-.0885	.0264	-.2457	93	.9193	.2510	-.0928	.0326	.1080
7	100	.8265	.3880	-.0074	.1363	-.1823	89	.8945	-.0458	-.2176	.1512	.1143
8	90	.1331	-.0343	.1308	.8614	.1178	79	.0566	-.0930	-.1463	.8471	-.2022
9	94	.5676	-.4943	-.0668	-.4653	.0509	79	.1632	-.6100	-.3758	-.4678	.1764
10	98	.7239	-.3536	-.2165	.0559	-.3627	82	.5537	-.5752	-.2954	-.1183	-.2914
11	96	.8308	.1840	.0637	.2443	.3429	91	.5649	-.1381	-.5634	.3612	.3453
12	95	.6464	-.2511	-.1481	.2155	.5577	86	.1308	-.3067	-.8082	.2478	.1857
13	89	.5730	-.3017	-.4987	-.0379	.3172	77	.1638	-.2458	-.8043	-.1674	-.0895
14	96	.6105	-.0649	-.6634	.0320	.2511	88	.3400	-.0202	-.8397	-.1521	-.1916
15	94	-.0058	-.4777	.7690	-.0037	.1534	84	-.3209	-.6301	.2635	.2607	.4536
16	94	-.0102	.4424	.4712	.5224	-.0966	70	.2449	.1924	.3638	.6718	.1392
17	99	-.2385	.8805	-.0719	-.1066	.2748	92	.1019	.8973	.0478	.0232	.3252
18	89	.2699	.2419	.3271	-.6046	.1522	63	.2164	-.0039	.0573	-.3553	.6714
19	93	-.0498	.7532	.1722	-.2780	.3674	81	.1298	.6463	.0337	-.0290	.6126
20	97	.7767	-.3495	.3779	-.1769	-.0765	91	.4427	-.7512	-.1377	-.0479	.3517
21	97	.7345	-.4315	.2986	.0765	.0398	82	.3309	-.7481	-.2753	.1641	.2248
22	94	.3627	-.8042	.1951	.0490	.1470	84	-.1706	-.8479	-.2909	.0705	.0525
E.Roots		7.4877	5.6581	2.3454	1.9070	1.3286						
% Trace		35.59	26.89	11.15	9.06	6.31	% Var.	26.36	24.52	14.17	9.06	11.02
Cum. % Tr		35.59	62.48	73.63	82.69	89.01	Cum. % Var.	26.36	50.88	55.05	64.11	75.12

Total Trace = 21.0401

95.64% of total variation is common.

* Communality

Table 6a. Relationships between Factors of the Light and Heavy-seeded lines (Correlations.)

Light-seeded	Heavy-seeded 1	2	3	4	5
1	.9570	.0551	-.1064	.1785	.1946
2	-.1801	-.2833	-.7882	.5115	.0660
3	-.0311	-.2918	-.2381	-.6274	.6808
4	-.2109	-.7353	.0184	.2869	.5764
5	.0784	.5394	-.5571	-.4801	-.4025

Table 6b. Relationships between identical pairs of variables from Analyses of Light and Heavy Sets.

Traits	Correlations
1	.5767
2	.8381
3	.5457
4	.9201
5	.9515
6	.9067
7	.9873
8	.9895
9	.6562
10	.9114
11	.9166
12	.9559
13	.6100
14	.5250
15	.5684
16	.7459
17	.7276
18	.1040
19	.2181
20	.8379
21	.5640
22	.8233

light-seeded results (Table 6 a-b, page 29). This confirms the finding that the first factor of the complete-set analysis is indeed a weight factor, since in none of the seed-weight sub-set analyses was it possible to recover it. These comparisons also indicate that the fourth factors of the medium-weight and heavy-seeded groups are almost the same. Furthermore, combinations of factors in one set may probably have the same effect of another factor or combination of factors in another set, an exception being made of the factors similar to the second factor of the complete set results (Table 7 a-b, 8 a-b, page 31).

These results were verified through multiple discriminant analysis. Of the two discriminant functions necessary to separate the three groups, both with very significant chi-squares, the first was the most important. It accounted for 88.60% of the variance (Table 8c, page 32). Seed weight (9), and the pod-dimension variables had the highest positive correlations with that function. This function is almost identical to the first factor-variable found through factor analysis of the complete set of data. The second function, on the other hand, resembles more the third factor from that structure. It is also similar, in pattern, but not in size of the loadings, to the discriminant function between vine and bush types, that is, it has high correlations with number of nodes per plant (1) and number of long internodes (17). These results show also that the three seed-weight groups can include branchy types of plants.

Table 7a. Relationships Between Factors Of Light-Seeded and Medium-Weight Lines (Correlations)

Light-seeded	Medium-weight 1	2	3	4
1	-.9372	.9933	-.0607	.0755
2	.4104	-.0138	.5254	.7388
3	.2505	.0710	.6005	-.6218
4	-.6779	-.0764	.0543	.2456
5	-.5549	.0487	.5973	-.0394

Table 7b. Relationships Between Identical Pairs of Variables From Analyses of Light-Seeded and Medium-Weight Lines

Variables	Correlations
1	.2834
2	.8427
3	.7824
4	.9850
5	.9880
6	.9089
7	.9409
8	.9716
9	.4034
10	.8408
11	.9861
12	.8770
13	.5852
14	.8363
15	.9276
16	.3917
17	.5085
18	.4928
19	.7206
20	.6114
21	.5232
22	.8319

Table 8a. Relationships Between Factors of Heavy-Seeded And Medium-Weight Lines (Correlations)

Heavy-Seeded	Medium-Weight 1	2	3	4
1	-.0008	.9810	-.0832	-.1731
2	-.9846	-.0037	.0212	.0024
3	-.1270	-.0632	-.7595	-.0958
4	.0557	.1002	-.4176	.8273
5	-.1067	.1538	.4913	.5257

Table 8b. Relationships Between Identical Pairs Of Variables From Analyses Of Heavy-Seeded And Medium-Weight Lines

Variables	Correlations
1	.9336
2	.8101
3	.7848
4	.9381
5	.9310
6	.9665
7	.9310
8	.7345
9	.9011
10	.7565
11	.9103
12	.7828
13	.5524
14	.5858
15	.6470
16	.9092
17	.9521
18	.4508
19	.8728
20	.9237
21	.9982
22	.9106

Table 8c. Multiple Discriminant Analysis between Light-seeded, Medium-weight and Heavy-seeded Lines

Variables	Discriminant Axes Loadings 1	2
1	-.2864	.5705
2	.0493	-.0356
3	.1351	-.1261
4	.0439	.0145
5	-.0634	-.0462
6	-.2134	.1899
7	-.4005	.1214
8	-.5265	.2964
9	.9115	-.1776
10	-.0127	-.1501
11	.4259	.0305
12	.8874	-.0830
13	.9159	-.1644
14	.8987	-.1240
15	.7612	.0784
16	-.2884	.1546
17	-.3580	.5227
18	.6435	-.3102
19	-11154	.1257
20	.5787	-.1801
21	.5693	-.0167
22	.5543	-.1267
% Of Total Variance :	88.60	11 .40
Chi-Square value :	292.655	105.486
Degrees of Freedom :	23	21
Significance Level :	.0000	.0000

DISCUSSION

Interpretation of the results

Factor analysis of the complete-set of data reveals that there are two major factors or patterns of production in beans and a much less important factor more related to the morphology of the plant than to its productivity. Thus, the relation of this third factor to production is indirect. The three factors have been tentatively called pod-weight, pod-number and growth, respectively. All three factors remained invariant from one location to another despite the different types of planting which were used to accentuate the differences between the two environments. It is to be noted, however, ~~that~~ that the means of many of the variables were found to be significantly different between locations.

Factor analyses with the growth types show that these two major factors exist among both determinate and indeterminate lines and furthermore, that whereas the first factor is the most important for the bush type, the second is probably the main factor for the vine type. The third factor was also recovered in both cases, but with a different configuration among the indeterminate lines. Consequently, it can be concluded that these patterns are consistent across locations and growth types.

Multiple discriminant analysis, both between growth types and between seed-weight groups, establishes beyond any doubt that the third factor of the complete-set results is indeed a growth factor. It was

responsible for about 10% of the total variation before rotation and 15% after rotation in the factor analysis of the complete-set of data. This indicate that growth habit has a relatively low impact on bean productivity. The factor of growth can be expressed in terms of the total number of nodes and the number of long internodes on the plant. Probably, the greater the number of internodes of the plant (presumably indeterminate), the thinner the stem and the pods and also the lighter in weight will be the whole plant. In an extreme case, the plant will certainly produce too many racemes, many undeveloped pods giving way to an unbalanced set of yield components, and consequently a very low yield.

The results of the factor analyses of seed-weight groups data are in accord with those of the complete-set. The first factor in the latter was in fact a weight factor, since it was not recovered in any of the seed-weight groups analyses. But,the second factor, or number-factor was identified in all of them. This indicates that the grouping of bean varieties by seed weight is very appropriate and its implication on productive potential is great, given the importance of the weight-factor in this study.

As a multiple discriminant function, the weight factor accounted for about 89% of the total variance. In the results from factor analysis of the complete-set of data, it was as important as the number-factor. This number-factor was recovered in all seed-weight group analyses and as in the indeterminate sub-set, it was the most important factor in the light-seeded sub-set results.

Implications of the results on the Development of an Improved Plant-type in beans

The results indicate that in order to reach a high yield through high seed weight, a plant has to have a sturdy stem and relatively few but long basal internodes. Such a plant can not therefore also have many nodes and certainly will have no or very few long internodes. It will be more feasible to increase the number of pods on that plant than to augment its number of seeds per pod. Its production will be mainly or uniquely located on the stem. The height of the plant is due mostly to the length of its basal internodes.

By contrast to this pod-weight type, in order to attain a high yield per plant through a high pod number, a bean plant needs many basal internodes which should be moderately sturdy. The number of nodes on the plant can be higher than in the preceding type. Its pod-bearing potential is higher on the branches than on the stem. This plant-type can have considerably more seeds per pod, but the seed-weight should be kept at a minimum. Interestingly, this pod-number type can weigh more than the pod-weight type and is also more efficient since its number of pods per productive node is higher. The height of the plant is due primarily to the number of basal internodes.

The factor related to growth shows that a bean plant with numerous, long and thin internodes will be an extremely poor yielder. This also will be a strong competitor. Consequently, in improving yield, either on a per-plant basis or on an area basis, such a plant-type should be avoided. However, this result can also be seen as an invitation for inter-breeding between the pod-weight and the pod-number

types, so long as it is possible to maintain a favorable balance between the number of nodes and the diameter of the stem. Whenever this can be done and at the same time the number of upper-internodes can be kept at a minimum, the yield of the plant, whatever its original plant-type will be improved considerably.

It is interesting to note that although most of the unadapted, south-american varieties were not included in the factor analyses, their yields and morphological characteristics strongly confirm the result that too many nodes or internodes on the plant, particularly the upper ones, will provoke a reduction in its yield,

It seems therefore feasible to attempt to develop an improved bean plant-type by increasing the pod-bearing capacity of the stem of a heavy-seeded line. This can be done by augmenting the number of its short or basal internodes and also maintaining the sturdiness of its stem. If the seed size can be kept at a level commercially acceptable, without increasing the number of branches, this superior plant-type will fit well to the objective of improving yield on an area basis, because inter-plant competition will be thus minimum. However, the pod-bearing capacity of the plant as a whole can still be enhanced by augmenting the number of its branches. This can be realized through back-crossing with a pod-number type parent. This may not appeal to most plant breeders in view of the fact that the inter-plant competition might be consequently increased. But, it still may be wise to do it, at least as an intermediate step, since this is likely to produce a reduction in seed size, something which in itself may be economically beneficial. Moreover, the pod-number type parent may also be used as

as the carrier of other useful traits. But, if the objective is to improve yield on a per-plant basis, the number of branches on the plant should always be a matter of concern, since the higher their number, the higher the yield. In fact, everything else being comparable, any bean plant will yield more with than without branches.

So, keeping in mind the limit imposed by the third factor, the growth factor, several improved bean plant-types can be obtained through repeated inter-breeding between the two major plant-types, namely the pod-weight and the pod-number types, Shifting emphasis, through increasing the number of backcrosses toward one type or the other, permits the plant breeder to enhance the differences between those improved plant-types, that is, to separate genetically the ones approaching his ideal of a good variety on a plant-yield basis from those getting close to his"ideotype" on an area-yield basis.

SUMMARY AND CONCLUSION

There results of this study indicate that there three uncorrelated factors or patterns of production in beans and that only the pod of the plant can be used to distinguish between the first two. The first most important factor is pod-weight, the weight of the pod as expressed through the size of both the seed and the pod. The second most important factor is pod-number and is related more with the number of racemes on the plant than with the number of pods per productive node. It is interesting to note that the number of branches on the plant has only a very low weight in the obtaining of this factor. Surprisingly, number of seeds per pod does not have a high positive loading on any of these two major factors. On the contrary, that variable has a negative and moderately high loading only in the first factor. The third and least important factor is a growth factor with emphasis on the general size of the plant. Incidentally, this factor will not necessarily discriminate between determinate and indeterminate lines. Its resemblance with the discriminant function calculated for that purpose indicates rather that the indeterminate lines are more likely to have many nodes and long internodes as compared to the determinate lines.

It was found that it is possible to make progress in yield improvement by breeding within any of the two major types, that is, increasing the weight of the seed of a pod-weight type variety or the number of pods of a pod-number type variety will often result in better yields. In the first case, the short or basal internodes

should be made longer and sturdier whereas in the second they should become more numerous while remaining sturdy. However, greater advance is obtainable through interbreeding between the two types. For example, the number of short internodes can be increased in a pod-weight type line so that it can produce more pods. In the same manner, the weight of the pod and the seed can be improved in a pod-number type variety by making its short internodes longer and sturdier. However, two variables should be kept under control in either case, the number of long internodes and the number of branches. The first will prevent the plant from approximating its maximum yield potential, and the second will make the plant a strong competitor, thus limiting its use to production on a per-plant basis.

Obviously, since beans are categorized for commercial usage largely on the basis of seed size, the plant breeder therefore must place greater emphasis upon obtaining a plant-type which, on an area basis, maximizes the pod number.

BIBLIOGRAPHY

BIBLIOGRAPHY

Bailey, D. W. 1956. A comparison of genetic and environmental principal components of morphogenesis in mice. Growth 20:63-74.

Cattell, R. B. 1965. Factor analysis: An introduction to essentials. I. The purpose and underlying models. Biometrics 21:190-210.

Harman, H. H. 1967. Modern factor analysis. The Univ. of Chicago Press 2nd Ed. Chicago. 474 p.

____________. 1968. Factor analysis. In. Handbook of measurement and assessment in behavioral sciences. D. K. Whitla ed. Boston: Addison-Wesley. pp. 143-170.

Hashigushi, S. and Morishima, H. 1969. Estimation of genetic contribution of principal components to individual variates concerned. Biometrics 25:9-15.

Hegmann, J. P. and DeFries, J. C. 1970. Maximum variance linear combinations from phenotypic, genetic and environmental covariance matrices. Mult.Behav. Research 5:9-18.

Morishima, H., Oka, H. and Chang, T. T. 1967. Analysis of genetic variations in plant-type of rice. 1. Estimation of indices showing genetic plant-types and their correlations with yielding capacity in a segregating population. Japan. J. of Pl. Breed. 17:1-12.

Murty, B. R. and Arunachalam, V. 1967. Factor analysis of diversity in the genus of Sorghum. Ind. J. of Genet. 27:123-135.

___________. and ______________. and Jain, O. P. 1970. Factor analysis in relation to breeding system. Genetics 41:179-189.

Rummel, R. J. 1970. Applied factor analysis. Northwestern Univ. Press. Evanston. 617 p.

Veldman, D. J. 1967. Fortran programming for the behavioral sciences. Holt, Rinehart and Winston. New York. 406 p.

MICHIGAN STATE UNIV. LIBRARIES
31293008239778