THE HEXAMETRICAL MAZE

Once upon a time it was perfectly easy to tell the difference between poetry and prote : if it rhymed or scanned it was poetry, otherwise it was prose. To be sure, as early as Aristotle (*Poetics* 1447 b 18-20) there is a demand that the subject matter be poetic, and Horace can deny (*Sat.* I, 4, 39 ff.) that what he is writing is poetry, but for a very long time it seems never to have been questioned that formal constraints such as meter or rhyme were necessary if not sufficient conditions for poetry. That that time is no longer with us is obvious, and we are all familiar with the arguments that brought about the rebellion against such constraints. There remains merely a lurking, naive suspicion that writing under formal constraint is prettier, more difficult, and hence better, but that notion is quickly dispelled for students in their earliest academic exposure to poetry, and probably rightly so, although we are not yet united on the idea that formal constraint is positively unpoetic.

In what follows, I wish to exploit this notion of constraint in an analysis of Latin dactylic hexameter, for I shall compare the formal constraints of the hexameter line to a kind of maze which must be repeatedly traversed by the poet. In so doing, I hope to throw some light on how some poets did or did not resemble each other in their metrical practices, and I shall attempt to show how these divergences may be measured.

To be sure, it may be argued that a constraint which is unquestioned is no constraint at all. One has the feeling that Ovid felt no difficulty in producing hexameter verses, and it may even be held that, under certain circumstances, a metrical scheme may serve as a crutch rather than a stumbling block. Indeed one may believe that some of the more arid stretches of epic are poetically sustained by nothing but their meter. Be that as it may, there can be no doubt that the dactylic hexameter imposed a discipline of sorts upon the poets to which they responded in characteristically different ways. Some modes of response may have been deliberate other the result of unconscious habit; it is difficult to make the distinction.

Let us begin with a prelude where the hypothesis of conscious choice seems intuitively most acceptable. I have compiled statistics on eight little words, monosyllables, which can be made long by position or left short at the discretion of the poet and which occur with reasonable frequency in all the works concerned. See Table 1.

TABLE 1

Total occurrences / Number long (Percentage long/Percentage ranking)

	Eclogues	*Aeneid I*	*Metamorph. XII*	*Ars Poet.*	*Culex*	*Total*
ab	18/1 (.06/8)	12/2 (.17/8)	11/0 (.00/8)	4/0 (.00/8)	9/1 (.11/8)	54/4 (.07/8)
ad	20/15 (.75/4)	21/17 (.81/4)	11/3 (.27/7)	13/4 (.31/7)	15/12 (.80/3)	80/51 (.64/5)
et	131/100 (.76/3)	161/141 (/85/2)	90/46 (.51/5)	144/95 (.66+/4)	81/63 (.78/4)	611/445 (.73/4)
in	66/41 (.62/6)	48/36 (.75/6)	76/21 (.27+/6)	35/23 (.66/5)	35/13 (.37/7)	260/144 (.55/7)
nec	52/52 (1.0/1)	18/18 (1.0/1)	26/20 (.78/3)	22/22 (1.0/1.5)	16/15 (.94/2)	134/127 (.95/1)
sed	13/11 (.85/2)	11/9 (.82/3)	16/11 (.69/4)	6/4 (.67/3)	7/7 (1.0/1)	53/42 (.79/3)
ut	29/21 (.72/5)	10/8 (.80/5)	15/12 (.80/1.5)	24/24 (1.0/1.5)	11/6 (.55/5)	89/71 (.80/2)
Total	336/244 (.73)	311/246 (.79)	250/117 (.47)	253/175 (.69)	190/125 (.66)	1340/907 (/.68)

Perhaps the most striking fact to emerge from the table is the degree to which the Ovidian selection differs from the rest in keeping these monosyllables short, and it seems clear that Ovid has deliberately cultivated a profusion of dactyls rather than spondees. It is no surprise to find the *Aeneid* more spondaic in tone than the *Eclogues*, but the difference between the difference between the two is reassuringly small. It is also interesting to note that the *Culex*, otherwise so close to Vergil in its characteristics, is quite strikingly different in its use of the word *in*.

The facet of the table with which we are especially concerned, however, is the relative ranking within works of the degree to which a word is made long, and it is here that the notion of deliberation seems less acceptable. Aside from the special cases of *ab* and *nec*, there seems to be no particular reason why any word should be lenghtened more frequently than any other, and the poems do differ in this regard to a degree, It is as if these words embody different degrees of length or emphasis for these poets despite their metrical equivalence. Certainly Ovid and Horace seem to feel that *ad* should generally be short, but this feeling is not shared by Vergil and *Culex*. Again, Ovid and *Culex* share a feeling about *in* not seen in the others, but this latter conclusion may be deceptive, since *in* is ranked 5, 6 or 7 by all the poets, and hence the disparity may rather be due to a shared tendency to avoid spondees. Finally, Horace seems convinced, at least here, that *ut* should be long.

The case of *ab* is different since the word tends to be used in front of initial vowels. *Nec* is complicated since it alternates with *neque*; indeed short *nec* and *neque* elided before an initial vowel are metrically indistinguishable, and our figures here may be especially prone to editorial whim.

(In the editions which I have used, elided *neque* occurs 8 times in the *Eclogues*, 3 times in *Aeneid I*, and once in Ovid, and not at all in the other two. Conflating the figures would not change the rankings.)

The similarity or disparity of the rankings may be described most succinctly by the rank-correlation coefficient, described in any text-book on statistics.* The ten coefficients which may be derived from these figures are given in Table 2. The coefficient itself is designated by r_s, and it is a value with theoretical limits of +1 and -1, the sign indicating positive or negative correlation. The statistic has inferential meaning as well, and @ designates the level of significance. Thus to take the outstanding case of the correlation between the *Eclogues* and *Aeneid I*, the coefficient is +0.98 and the level of significance is less than .001. The latter figure means that out of all the possible different comparative rankings (in this case there are 8! x 8! or over 1.5 billion), less than one in a thousand will have a coefficient of rank-correlation so close to +1. We must, however, not be overly impressed by these numbers. As is usually the case in inferential statistics, the null hypothesis to be accepted or rejected must itself be able to withstand the rigors on an *a priori* analysis. In this case, our only null hypothesis can be that the coefficient is the result of chance. We must also select a level of significance, which means, simply, that we must define the amount of risk of being wrong that we are willing to accept. The selection of such levels is a matter of both statistical convention and experience. For example, much scientific work is carried on at the 5% level, in which case the null hypothesis would be rejected, in the first 6 cases in Table 2.

---

* See, e.g. W. J. DIXON and F. S. MASSEY, *Introduction to statistical Analysis* (McGraw-Hill, 1969^3) pp. 349 ff.

TABLE 2

	r_s	@
Eclogues - Aeneid I	0.98	$<$ .001
Eclogues - Culex	0.93	.001
Aeneid I - Culex	0.86	.005
Eclogues - Ars Poetica	0.69	.035
Metamorph. XII - Ars Poetica	0.69	.035
Aeneid I - Ars Poetica	0.67	.042
Ars Poetica - Culex	0.60	.066
Metamorph. XII - Culex	0.31	.231
Eclogues - Metamorph. XII	0.24	.291
Aeneid I - Metamorph. XII	0.21	.310

But let us say, for the purposes of discussion, that the null hypothesis has been rejected at a reasonable level of significance. We have concluded that the coefficient is not the result of chance. Of what, then, is it the result ? No immediate answer is forthcoming. The number of potential null hypotheses is unlimited and we can only go on, if possible, rejecting them one by one, providing they have been properly formulated. However, the picture is not quite so hopeless as it appears. For while we can do no more with the present case, we have learned something for next time, although not very much. Lurking in the background is the question of authorship. Based on our present slender experience, and due to the fact that we have used "cooked" material (i.e., we know that the *Eclogues* and *Aeneid I* have a single author, and that the *Eclogues* and *Ars Poetica* do not), we may with some slender justification formulate the following null hypothesis for a different case where the same test would be used : the two texts are by different authors; this to be rejected only at a level of significance less than .001. If we were convinced that the *Culex* is by Vergil, we might set the level of significance at less than .035. This, then, would be a case of arriving at a proper level of significance through statistical experience, but it need hardly be stated that an empirical induction of this sort, resting as it does on the single case, is a slender reed indeed. As in all cases of induction, we simply need much more experience before we can proceed with confidence. What we do know, at least, is that a level of .035 or greater is not sufficient to differentiate between authors so far as this test is concerned.

So much for the inferential aspects. For the moment, it seems preferable to be satisfied with description. We have seen how and to what degree the various works resemble each other in the use of these monosyllables. Using the rigor of, say, an historian rather than a statistician, we may conclude that Ovid used these easily manipulable words so as to emphasize the dactyllic aspect. To a lesser degree, Vergil uses them in the *Aeneid* so as to emphasize the spondaic. It is, however, difficult to believe that the consistency of ranking in Vergil was due to conscious policy, and there is enough evidence at the 5% level of significance, to conclude that all the authors involved shared a general feeling about the comparative degree to which these monosyllables should be lengthened.

Having done with this brief prior analysis of one of the more obvious means by which poets could ease their way through the hexameter, we must now turn to our major task which is the comparison of the hexameter to a complex maze. Our major hypothesis is quite simple : faced repeatedly by a complex task, individuals (both poets and, with immediate apologies to the Muse, white rats) tend to develope characteristic and individual ways of accomplishing that task. In Skinnerian terms, repeated success will tend to reinforce those ways that work, and, in a reasonable application of the principle of the conservation of energy, the poet (and the rat) will develope habitual modes of behavior, that is, modes of behavior which do not require conscious thought, laborious decisions, and the constant exploration of alternatives. We need not debase ourselves excessively; so far as I know, no white rat has ever mastered a task as complicated as the writing of a hexameter verse, and to paraphrase Aristotle, man is the animal that writes hexameters.

The hexameter line contains 12-17 syllables, and may be described in its 17-syllable form as follows :

0 — 1 ∪ 2 ∪ 3 — 4 ∪ 5 ∪ 6 — 7 ∪ 8 ∪ 9 — 10 ∪ 11 ∪ 12 — 13 ∪ 14 ∪ 15 — 16 ∪ 17

Numbered from 0 - 17 are the 18 points in the line where word-ending and/or word-beginning can occur. In the notation which is adopted here, the substitution of spondees for dactyls simply means that no word-juncture occurs at points 2, 5, 8, 11 and/or 14 depending on which and how many feet are made spondaïc.

We may begin by listing the percentage of verses having word-juncture at each of these 18 points. The three works with which we shall be concerned are the *Eclogues*, *Metamorphoses XII*, and the *Culex*. The percentages and rankings are listed in *Table 3*. If we did not already know it, a glance at *Table 3* is sufficient to convince us that simply combining a set of words to fit the meter will not do. It is true that word-juncture can occur at any point in the line, but it is also clear that some points are favored far more than others, and that all three works are in substantial agreement on which points these are and the order in which they are to be prefered. The rank-correlation coefficient in each case is +.98 with a level of significance lower than .001. On the basis of these poems at least, one is not traversing the maze properly, i.e., not writing proper hexameter unless about 85% of the verses have word-juncture at point 7 (the main caesura), about 65% at point 10, and so on, while at the same time, of course, respecting the meter.

What we have seen thus far servers better to show how these poems are similar rather than different, and we may justly suspect that what we have here is approximately true of all hexameter poems. There are some interesting divergences in Table 3. For example, the "bucolic diaeresis" at point 12 is more common in the *Eclogues* and *Culex* than it is in *Metamorphoses XII*. The largest disparities occur at point 1 between the *Eclogues* and *Metamorphoses* and at point 4 between the *Eclogues* and *Culex*. More analysis is necessary however, before proceeding along these lines, if we wish to arrive at a legitimate measure of the differences between these poems with regard to word-juncture. The major theoretical problem is that we do not wish to weigh the same evidence twice.

TABLE 3

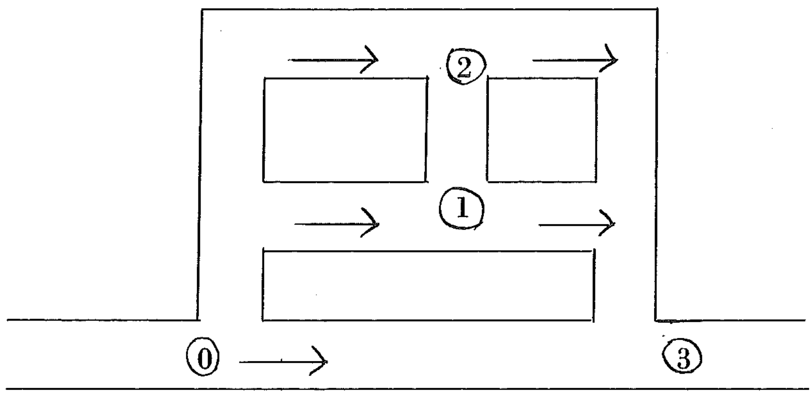
Percentage and (ranking) of verses with word-juncture at particular points

Points	Eclogues	Metamorph. XII	Culex
0	100.0	100.0	100.0
1	40.4 (8)	28.2 (8)	30.4 (8)
2	18.6 (10)	25.2 (9)	16.5 (10)
3	51.1 (6)	52.6 (6)	45.5 (6)
4	64.2 (4)	56.5 (4)	52.6 (5)
5	17.8 (11)	15.8 (11)	17.0 (9)
6	10.5 (13)	4.6 (14)	6.1 (13)
7	85.5 (1)	89.1 (1)	84.9 (1)
8	13.9 (12)	12.0 (12)	11.9 (12)
9	21.3 (9)	21.4 (10)	12.6 (11)
10	71.4 (2)	64.6 (2)	66.7 (2)
11	4.0 (15)	6.9 (13)	4.6 (15)
12	64.3 (3)	54.5 (5)	63.8 (4)
13	4.4 (14)	1.1 (15)	5.1 (14)
14	41.5 (7)	45.0 (7)	41.1 (7)
15	60.3 (5)	63.8 (3)	65.7 (3)
16	1.6 (16)	1.0 (16)	2.0 (16)
17	100.0	100.0	100.0

For example, we see that the *Eclogues* have word-juncture at point 1 in 40% of their verses. *Metamorphoses XII* has such word-juncture in only 28% of its verses. This difference of 12% is sizable and certainly noteworthy. Taken as it stands, it is a perfectly legitimate observation of a difference between these two works. If, however, we go on to say that the works also differ at point 2, since the *Eclogues* have word-juncture here only 19% of the time, while *Metamorphoses XII* has it 25% of the time, than we must be careful. This second observation is just as legitimate as the first, but we must not go on to say that we now have two indications of the differences between the two works, for while we have two indications, they are surely not two *independent* indications. Both poets must begin their line with a word, and any word which does not terminate at point 1 must terminate elsewhere. If fewer of Ovid's initial words terminate at position 1, then we must not be surprised if more of them terminate at point 2, and it would be delusive to say that we have two independent measures of difference. We are surely describing the same difference twice. It seems more reasonable to observe that there is also a difference of 10% at point 12 and, perhaps, to hope that the observations at points 1 and 12 are independent. For while it is clear that the observations at adjacent or nearly points are linked, one may surmise that the aspect of dependence has dissipated in the distance between points 1 and 12, and that perhaps the poet's behavior at point 1 does not affect his behavior at point 12.

As a result, given the dependence linkage effect described above, the information in Table 3 cannot be fully utilized. We would be justified in using the observations at a single point only, although we might hope that the observations at one or two other widely separated points would also be substantially independent.

Happily there is a better mode of procedure, but one which requires more information than that given in Table 3. We shall employ descriptively an elementary bit of Markov chain analysis, omitting all the theoretical implications. Such analysis involves a transition-matrix of probabilities, and it is this that we shall construct. Let us think now of white rats as they move through a maze with four points in it numbered from 0 to 3. The maze is so constructed that the rats can go from any lower-numbered point to any higher-numbered one; but not the reverse. Such a maze might look like the following :



Let us observe a hypothetical white rat (A) as he goes through this maze 100 times. Our observations may be conveniently noted in the following matrix. At the left are the points which he travels from. Across the top are the points to which he goes directly.

	1	2	3	
0	20	20	60	
1		5	15	Rat (A) actual
2			25	

Thus me see that our rat (A) goes directly from point 0 to point 3 sixty times, from point 1 to point 2 five times, and so on. If, on the basis of our 100 observations, we wished to predict, as best we could, the path of rat (A) on the next trial, we would do so by converting the matrix to percentages based on the horizontal row sums as follows :

	1	2	3	
0	20%	20%	60%	
1		25%	75%	Rat (A) percentages
2			100%	

The above is a transition-matrix of probabilities, and it is the basis for Markov chain analysis. For example, we can calculate the probability of a path through all four points by finding the product of the probabilities of each part of the path : from 0 to 1, 20%; from 1 to 2, 25%; from 2 to 3, 100%. The product is 5%, so there is one chance in twenty that the rat will take the prescribed path. (I omit here all considerations of standard error.)

For purposes of furthering the discussion, let us now observe hypothetical rat (B), whose performance on 100 trials is recorded as follows :

	1	2	3
0	40	40	20
1		10	30
2			50

Rat (B) actual

At first glance, we may be tempted to say that rat (B)'s performance differs substantially from that of rat (A) throughout the maze, just as we have previously said that Ovid differs from Vergil at both points 1 and 2. But if we look at the transitionmatrix of probabilities for rat (B) :

	1	2	3	
0	40%	40%	20%	
1		25%	75%	Rat (B) percentage
2			100%	

we see that, although the rats differ markedly in their behavior upon leaving point 0, they behave exactly the same upon leaving point 1. However, as a result of their different behavior at point 0, rat (B) reaches point 1 twice as often as rat (A), and their similarity of behavior at point 1 is masked until revealed by this sort of analysis. In exactly analogous fashion, we may say that Vergil's habit of employing initial monosyllables guarantees his having word-juncture at point 1 more frequently than Ovid, but employing this mode of analysis, we may go on to argue as follows : given the fact that Ovid has word-juncture at point 1 less often than Vergil, it is still possible to make an independent observation of what both poets subsequently do in all those cases where word-juncture does in fact occur at point 1. Their subsequent behavior at this point may or may not differ. What is important is that our present observation is independent of the former and that we can make such independent observations at every point in the maze or line. As a result, we can use all of our information, rather than startling bits of it, to arrive at a summary judgment of similarity or difference in behavior. We thus avoid both the difficulty of weighing the same evidence twice or even more frequently and the problem of the subjective selection of evidence.

In the above discussion, it has been tacitly assumed that writing a hexameter verse is much like moving forward through a maze, i.e., one writes the first word, then the second, and so on, but, of course, this cannot be the case. There are, and always were, innumerable ways of composing hexameters, all the way from the perfectly formed verse springing forth full-blown to the laboriously backed exercise of the schoolboy. No one, except perhaps the poet, can say where or how a verse begins to be formed, and the question would be otiose and irrelevant to my discussion, were it not for the following point : the transition-matrix of probabilities is changed rather drastically if the direction of travail is reversed. Let us return quickly to hypothetical rats (A) and (B), but now imagine that they have travelled the same routes in the reverse direction :

	Rat (A) actual reverse (percent)			Rat (B) actual reverse (percent)		
	0	1	2	0	1	2
1	20 (100)			40 (100)		
2	20 (80)	5 (20)		40 (80)	10 (20)	
3	60 (60)	15 (15)	25 (25)	20 (20)	30 (30)	50 (50)

Once again, our rats differ only in their behavior at the initial point of the maze, but where previously this contrasting behavior was represented by the probability vectors (20, 20, 60) and (40, 40, 20), it is now represented by the probability vectors (60, 15, 25) and (20, 30, 50), (a vector is a row or column in a matrix) and these contrasts are not at all alike. It is difficult to imagine the incompetence of the observer who does not know the direction in which his rats have traversed the maze, but that is exactly our situation with regard to the hexameter. All that our observations tell us is that, e.g., *Metamorphoses XII* has 37 words that extend from point 0 to point 5, and 105 words that extend from point 5 to point 7. We do not know which word was "composed" first, and, in a way, it does not even make sense to ask such a question. As a result, we are left in some uncertainly, since the analysis will differ depending on whether we inspect the verse moving forward from the beginning or backward from the ending. With the aid of the computer, however, we can do the analysis in both directions and compare the findings. As we shall see, this comparison will allow us in a rather surprising way to state after a fashion which word was "composed" first. Table 4 contains the lengths and positions of words in *Metamorphoses XII* in the same matrix from used to describe movements in the rats'maze.

By using the horizontal row totals in *Table 4,* we could construct a transition-matrix of probabilities for movement forward through the verse. Just so, the vertical column totals could be used to construct a similar matrix for backward movement. We might construct similar matrices for the *Eclogues* and *Culex*, and then compare them. However, Table 4 is presented only for illustration, and we shall now proceed to adopt a somewhat different strategy for the following reasons :

TABLE 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0	177	152	168	91	37		2										
1		6	106	53	11		1										
2			50	91	10		7										
3				122	45	4	153										
4					4	27	280	46									
5						2	105										
6							19	13		1							
7								21	135	274	28	104		5			
8									6	66		3	1	4			
9										67	8	48		13	5		
10											8	180		138	82		
11												7		9	28		
12													4	114	222		2
13															4		1
14															62	7	214
15																7	395
16																	14

First and foremost, we wish to make a statistical comparison of the matrical behavior of three poems, but I do not know how to make such a comparison of transition-matrices of probability. We are dealing with a rather bizarre problem, and I have not been able to find an account of such comparison that I can comprehend.

Second, Table 4 does not contain all the information which I intend to utilize, for it conflates a number of details which I shall keep distinct. Briefly, Table 4 does not distinguish between normal word-juncture, and juncture with elision or ecthlipsis. Also, I wish to distinguish between words embodying two short syllables as opposed to a single long in the second half of the foot, and Table 4 groups these together.

It would be possible to construct a matrix which contained all this information, but it would be much larger, and in any case, the first objection noted above would still hold.

The alternative adopted here is an extensive application of the chi-square test in a way which utilizes the analysis made thus far.

To clarify the discussion, we turn once again to the simple example of our two white rats. In moving forward through the maze, the rats differed in their choices at gate 0 as follows :

Choices	Rat (A)	Rat (B)
0 - 1	20	40
0 - 2	20	40
0 - 3	60	20
	100	100

The first question we may ask is whether these sets of choices differ to a significant degree. The alternative would be that the choices of the rats do not differ to a significant degree and the differences noted are simply due to chance variation. This last phrase should be taken in the following way : if we took 200 identical disks, wrote 0-1 on sixty of them, 0-2 on sixty, and 0-3 on eighty of them, mixed the disks thoroughly in a hat, and then divided them at random into two piles of 100 disks each, what would the likelihood be of getting a division like the one described by the rats' behavior ? We begin by making a best estimate of what the result of such a hat-experiment would be. This is exactly analogous to our best estimate of what the result of 100 tosses of a fair coin would be : 50 heads and 50 tails. We would, in fact, be very surprised if exactly 50 heads and 50 tails turned up, but we would expect the actual result to be closer to this estimate than any other that we could make. This best estimate is given the name "expected value". The expected values in our case are given below in parentheses :

	Rat (A)	Rat (B)	Row totals
0 - 1	20 (30)	40 (30)	60
0 - 2	20 (30)	40 (30)	60
0 - 3	60 (40)	20 (40)	80
Columns total	100	100	Grand total 200

Here, the expected values are the same for both rats because the column totals are the same. In general, the expected value for each cell is derived by dividing the product of the cell's row and column total by the grand total.

The partial chi-square value for each cell is produced by dividing the square of the difference between the observed (or actual) value and the expected value by the expected value. Thus, the partial chi-square value for the cell on the upper left is $(20\text{-}30)^2/30 = 3.33$. We repeat the tableau with the partial chi-square values in square brackets :

	Rat (A)	Rat (B)
0 - 1	20 (30) [3.33]	40 (30) [3.33]
0 - 2	20 (30) [3.33]	40 (30) [3.33]
0 - 3	60 (40) [10.00]	20 (40) [10.00]

A look at the partial chi-square values is reasonably informative, although not in the present simple case. They make clear that the rats differ most in their propensity to select 0 - 3.

The chi-square value for the entire tableau is the sum of the partial chi-square values for all the cells. In this case, chi-square (χ^2) = 33.32. We must now find a rather abstruse quantity known as the degrees of freedom (df). This is a difficult topic, but some rules of thumb are available. In our present case, we are holding the marginal totals (i.e., row totals and column totals) constant. As a result, the value of only two cells can vary (e.g., the first two cells in the left hand column). Once these two have been fixed, the values for all six cells are determined. In general, in a tableau of this sort, the number of degrees of freedom is found by multiplying the number of columns minus one by the number of rows minus one. In this case, df = 2.

If we now look in the proper statistical table (usually entitled "Percentiles of the χ^2 distributions") in the row for two degrees of freedom, we shall find that the highest value that occurs in this row is 15.20 and that it is in column labelled P 99.95. The conclusion to be drawn in the following : the hat-experiment we have described has a very large number of different possible results. Of all these possible different results, 99.95% will produce a chi-square value of 15.20 or less. In other words, if the experiment were repeated 10,000 times, we might expect to find a result as large as ours about 5 times. The level of significance (on the two-tailed test) is 0.001 which is as far as good tables go, and we may conclude rather confidently that the divergence in behavior between the two rats is not simply due to chance. Since these are imaginary rats, there is no point in further speculation about the reasons for this divergence of behavior; what is to be stressed is that we have here a measurement of that divergence, namely χ^2 (2 df) = 33.32. This value is comparable only with other chi-square values with two degrees of freedom, so we report the percentile value as well in the following eccentric fashion :

$$\chi^2 \text{ (2 df)} = 33.32 \text{ (99.95 - 100)}$$

indicating that the chi-square value lies between these percentile limits. The percentile value may be used to compare chi-square values with differing degrees of freedom. Further, it is most important to note that if we use the chi-square value, as here, to serve as a measurement rather than as a basis for inference, we do not use and to not need to use the words "random", "chance", "likely", or "probable". As said above, 99.95% of the possible different results will (not "would") produce a chi-square value of 15.20 or less. No inference is involved here. We are not opposed in principle to the utilization of statistical inference, but it seems desirable to make clear just when and how it is present. For if our purpose were purely descriptive, it would be possible to use means less arithmetically cumbersome than the chi-square statistic. The latter has, however, some striking virtues.*

Let us suppose that we give our rats another 50 trials each at point 0, and let us suppose that the results noted this time are rather different from those gained in the first set of 100 trials each. For example :

	Rat (A)	Rat (B)
0 - 1	10 (20) [5.00]	30 (20) [5.00]
0 - 2	20 (15) [1.67]	10 (15) [1.67]
0 - 3	20 (15) [1.67]	10 (15) [1.67]

$$\chi^2 \text{ (2df)} = 16.68 \text{ (99.95 - 100)}$$

There is no reason to believe that these results are any more or less valid than the previous ones for the purpose of testing the inferential null hypothesis that the divergence at point 0 is due to chance variation, and once having made these observations, it would be very poor practice to ignore them. One obvious and legitimate way of combining the results of the two sets of trials is to sum the figures for each cell and to create a new tableau, as follows :

* A very good presentation in S. Siegel, *Nonparametric Statistics : for the Behavioral Sciences*, (Mc Graw-Hill, 1956).

	Rat (A)	Rat (B)
0 - 1	20+10 (50) [8.00]	40+30 (50) [8.00]
0 - 2	20+20 (45) [0.55]	40+10 (45) [0.55]
0 - 3	60+20 (55) [11.36]	20+10 (55) [11.36]

$$\chi^2 \text{ (2df)} = 39.82 \ (99.95 - 100)$$

However, it is a remarkable fact about chi-square that it is also legitimate for inferential purposes to simply sum the chi-square values and degrees of difference of the 100-trial tableau and the 50-trial tableau, as follows :

$$\text{100-trial} : \chi^2 \text{ (2df)} = 33.32$$
$$\text{50-trial} : \chi^2 \text{ (2df)} = 16.68$$

$$\text{combined} : \chi^2 \text{ (4df)} = 50.00 \ (99.95 - 100)$$

Indeed, if concurrent divergence is a possible factor in the experimental design, this latter mode of combination is preferable.*

The above is a fair description of procedures involved in taking successive samples in order to test an inferential hypothesis. These samples are selected out of a reasonably well-defined parent universe or population which may be called the behavior (actual and potential) of rats (A) and (B) at point 0. By taking successive samples, we have arrived, presumably,

* The additive property of chi-square is discussed in E. Morice and F. Chartier, *Méthode Statistique : Deuxième partie : Analyse statistique* (Paris : Imprimerie Nationale, 1954) 246 ff.

at a better estimation of what the characteristics of that parent population are. Combining the chi-square values in the manner described above, is much like "averaging" the results of a series of tests in order to arrive at a fairer estimation of whatever it is that we are testing.

We are, however, interested in finding some single expression which will best characterize the total divergence in behavior at different points in the maze rather than sets of repeated trials at the same point.

From our previous discussion, it should be clear that a chi-square tableau may be constructed for each point in the maze where variation in behavior is found. For example, here is the tableau for our rats moving forward at point 1 :

	Rat (A)	Rat (B)
1 - 2	5 (5) [0]	10 (10) [0]
1 - 3	15 (15) [0]	30 (30) [0]
	20	40

$$\chi^2 \text{ (1 df)} = 0 \quad (0.0 - 0.5)^*$$

* Yates' correction for continuity in cases of one degree of freedom has not been applied here or subsequently. Its use would not substantially affect our findings. See Dixon and Massey, *Introduction to Statistical Analysis* (McGraw-Hill, Third Edition, 1969) 242.

The chi-square value reiterates, as expected, that there is no divergence of behavior hare, and we have a measurement which is independent of that found at point 0.*

In order to arrive at a summary expression of the divergence in behavior at both points 0 and 1, we propose simply to add the results together, just as if we had the results of successive trials at the same point. In this manner, we arrive at the following summation :

$$\begin{array}{ll} \text{point 0} & \chi^2\ (2\ \text{df}) = 33.32 \\ \text{point 1} & \chi^2\ (1\ \text{df}) = \ \ 0.0 \\ \hline & \chi^2\ (3\ \text{df}) = 33.32\ (99.95 - 100) \end{array}$$

This procedure is quite unusual, and it is not at all clear that we can have any confidence in the inferential aspects of this result. For point 0 is not identical with point 1, and it may be held that we are combining results derived from quite different criteria. From this point of view, it is much like combining the results of a test in maze-running and, let us say, fertility statistics for our two rats. In orthodox statistical procedure, what we would apply here, if anything, would be a test of correlation, i.e., does the more fertile rat chosse 0-1 significantly more often ? or something of that sort.

---

* From an inferential point of view, a percentile of 0.05 is just as unlikely as a percentile of 99.95. Yates' correction (see previous footnote) yields a percentile of 10-20, clearly more acceptable in the present case if the percentile level were to be inferentially applied.

On the other hand, it may be argued that one gate in a maze is very like another, and that we are taking successive samples from a parent population which may be tolerably well defined as behavior at a gate in a maze, or in the case of our poets, as behavior at a point of word juncture in a hexameter line. We are, therefore, in the obvious position of a teacher who tests his students with a series of geography tests, all different but similar, and then attempts to arrive at a summary grading of his students' geographical ability. In such a case, the teacher's best virtue is consistency, although there is always the further possibility of "standardising" the results of the battery of tests. Such a task, however, would involve the accumulation of far more data than we as yet possess.

In the light of this argument and counterargument, we shall sum our chi-square values for each point in the maze or line because the totals will prove useful for comparative and descriptive purposes, and we shall be consistent in arriving at these totals, but we shall look upon them as "raw" scores, and not attempt to use them for purposes of statistical inference.*

* I am not convinced that these totals are bereft of inferential value, but the theoretical difficulties are too large for me to solve. The problem, in brief, is the following : while we can blithely add chi-square values together, we cannot, in the present sort of case, construct a single tableau which will contain the combined eivdence as we could in the case of repeated sets of trials at the same gate. None of the other three approaches that have occurred to me seems preferable. One would be to construct (in the case of the rats) a 5x2 tableau incorporating the numbers of choices at both points 0 and 1, but this would destroy the independence of observation which our analysis has established by blurring concurrent divergences. Further, the total number of choices at point 1 is dictated by the number of times 0-1 is chosen, and a serious problem regarding the number of degrees of freedom arises, particularly in the complex case of the hexameter. A second approach would be a 5x2 tableau using percentages instead of numbers. This would not solve the problem of concurrent variation, and while it would preserve independence of observation, the inferential value of the chi-square test is based upon actual frequency rather than percentage. A third approach involving the number of different paths chosen seems impractical and is discussed in the text. I remain tolerably convinced that if any way of combining the results is acceptable, it is the one employed here.

It is worth noting that the chi-square score arrived at above (χ^2(3df) = 33.32) is identical with the result we should get rather more legitimately by a tableau incorporating the choices of the different paths through the maze rather than the choices at each gate. While this happy result is no more than a coincidence in our simple example, it suggests that an analysis according to different complete paths chosen would be an improvement over a dubious attempt to combine values achieved at each gate. There are, however, serious difficulties in applying such a mode of analysis to the complex maze of the hexameter line. Given the 18 points or gates in the hexameter, even if we do not consider abnormal word-juncture, the member of possible different lines is 2^{16}, or approximately 65,000. To be sure, a great number of these would not occur (e.g., there are no hexameters consisting of a single word), but given a work the size of the *Eclogues*, qe may expect that a large number of the lines will be unique. Uniqueness, however, is not a desirable statistical quality. Thus the textbooks tell us that proper utilization of chi-square analysis requires that none of the expected values should be less than one and that not more than 20% of them should be less than 5.* As a result, the data would necessarily be grouped, but it is not clear what orderly basis there would be for such grouping. On the other hand, it must be acknowledged that our present approach, despite its concentration upon behavior at each point in the line, will not reveal to what degree the verses of an author are metrically identical throughout their extent.**

* See, e.g., W.J. Dixon and F. J. Massey, *Introduction to Statistical Analysis* (McGraw-Hill, third edition, 1969) 238.

** Professor Stephen Waite has compiled statistics of exactly this sort for Statius' *Achilleid.*

Having at long length laid our theoretical foundation, we may at last devote our entire attention to the hexameter verse. We begin with automatic scansion* and the isolation of metrical word-types, using as an example *Metamorphoses XII,* 620 :

-	ipse	-	etiam	-	ut	-	cuius	-	fuerit	-	cognoscere	-	posses	-
-	2	&	1 1	+	2	-	2 2	-	11 2	-	2 2 1 1	-	2 1	
	1				2		3		4		5		6	

The above is the notation we have adopted for the mechanical scansion of Latin verse. The ampersand (&) indicates elision; the plus-sign (+) indicates ecthlipsis; the minus-sing (-) indicates normal word-juncture. Note that the scansion program makes no attempt to distinguish long from short in the final syllable of the line; they are all marked short.

The output of the scansion program was entered upon tape in the form above. It was necessary to develop methods of correcting the tape, since the scansion program does not scan all verses successfully. Once a complete text on tape was achieved, the computer was instructed to isolate each word with its attached scansion in the following way :

* For a discussion of automatic scansion, see my article in *Revue* 1967 (n° 3). The program has now been adapted for single-pass use on a computer of medium size. My colleague, Professor James Helm has revised the program so thoroughly that he deserves full credit for its present form.

ipse	et		-(1)2&
etiam	ut	(1)2	&11+
ut	cu		+(2)2-
cuius	fu		-2(3)2-
fuerit	co		-11(4)2-
cognoscere	po		-2(5)211-
posses			-(6)21

The beginning of a foot is indicated by a numeral within parentheses followed by the length of the syllable which initiates that foot. In cases like *etiam* where the foot is not indicated, a small adjustment in the program causes the computer to indicate the foot in which such scansions occur. In addition, the first two letters of the succeeding word are retained as useful information.

The computer was then instructed to construct two *indices verborum*, one with the words in alphabetical order together with their metrical word-types arranged in order under a secondary sorting. The second index was arranged according to a primary sorting of metrical word-types with a secondary sorting of the words in reverse alphabetical order. Concurrently with the production of this second index, the computer was instructed to punch a card for each metrical word-type indicating the work in question, the word-type, and the number of occurrences. These cards constituted the data-input for the statistical tests already described in principle. Given our previous discussion, the procedure may be described very briefly. In the first series of tests, the distributions of all metrical word-types sharing a common *initial* point and mode of word-juncture were compared. In the second series of tests, the criterion was a common *terminal* point and mode of word-juncture. Given the numerical restrictions of the chi-square test, the program was so constructed that only word-types occurring ten or more times in both works combined entered the calculation as individual categories. All others were grouped together in the bottom row of each tableau to become in themselves a single important statistical category.

TABLE 5

	Eclogues			Metamorphoses XII		
Metrical Word-type (M.W.T.)	Number	%	Partial chi-square	Number	%	Partial chi-square
(3)2 -1-	8	1.13	2.5395	17	3.04	3.2074
-11(4)2-	148	20.96	0.4473	103	18.43	0.5650
-11(4)21-	5	0.71	0.4301	7	1.25	0.5432
-11(4)22-	9	1.27	3.5809	21	3.76	4.5226
(3)2 -11-	39	5.52	0.7214	22	3.94	0.9111
-2(4)2&	8	1.13	4.1393	21	3.76	5.2278
-2(4)2-	226	32.01	3.1314	134	23.97	3.9549
-2(4)21-	15	2.12	0.8329	19	3.40	1.0519
-2(4)211-	34	4.82	0.0079	26	4.65	0.0100
-2(4)22-	75	10.62	0.3113	51	9.12	0.3932
(3)2 -2&-	9	1.27	0.0006	7	1.25	0.0007
(3)2 -2-	116	16.43	0.0987	98	17.53	0.1247
-11(4)2&-	0	0.00	6.6972	12	2.15	8.4584
All others (A.O.)	14	1.98	1.5676	21	3.76	1.9798
	706			559		

χ^2 (13 df) = 55.46 (99.95 < P < 100.0)

As an example of our detailed findings,* Table 5 displays in part (expected values and row totals have been omitted) the comparison between the *Eclogues* and *Metamorphoses XII* for words which begin with normal word-juncture at point 7, the main caesura of the line. The chi-square value for the divergence at this point is :

$$\chi^2 \text{ (13df)} = 55.46 \text{ (99.95 - 100)}$$

The partial chi-square values allow us to give an ordered description of the divergence between there works at this point. For instance, the five word-types contributing most of the total chi-square value in descending order are :

		Eclogues	Met. XII
	-11(4)2&-	0	12
	-2(4)2&-	8	21
	-11(4)22-	9	21
	-2(4)2-	226	134
(3)2	-1-	8	17

* It is neither possible nor desirable to reproduce here all the computer print-out generated by this project. As is usually the case, the equivalent of weeks' or months' sustained hand calculation was produced in seconds. For example, the comparisons of the *Culex* with other works required over 2,500 lines of printing and, among other tasks, the calculation of over a thousand partial chi-square values. The job required about 170,000 positions in memory and 67 seconds of real time on the IBM 360/50-75 at the University of Illinois.

Comparison of the same feature in the *Culex* and *Metamorphoses XII* gives the following result : χ^2 (13 df) = 61.34 (99.95-100). The divergence is, therefore, slightly greater than that found above. In this case, the five word-types contributing most are in descending order :

		Culex	Met. XII
	-2(4)2-	131 (37.54 %)	134 (23.97 %)
(3)2	-1-	0	17 (3.04 %)
(3)2	-2-	34 (9.74 %)	98 (17.53 %)
	-2(4)2&-	2 (0.57 %)	21 (3.76 %)
	-11(4)2&-	0	12 (2.15 %)

Turning to the *Culex* and the *Eclogues,* we find the following :
χ^2 (10 df) = 38.62 (99.95-100). The result is still significant, but the divergence between these two is less than that found in the previous two cases, and only two word-types make glaringly large contributions to the total chi-square value :

		Culex	Eclogues
	-11(4)22-	23 (6.59 %)	9 (1.27 %)
(3)2	-2-	34 (9.74 %)	116 (16.43 %)

Logically enough, these are the two word-types which were not schared in the previous two listings.

In exactly similar fashion, we may compare the three works with regard to metrical word-types which terminate at point 7 with normal word-juncture. The results follow :

Eclogues - Met. XII	χ^2 (9 df) =	39.95 (99.95 - 100.0)
Met. XII - Culex	χ^2 (10 df)=	20.76 (97.5 - 99.0)
Eclogues - Culex	χ^2 (9 df) =	34.60 (99.95 - 100.0)

We may, perhaps, find these results surprising. The works which previously shared the largest divergence now share the smallest. (For the curious, the *Eclogues* use a monosyllable at this point three times as often as the *Culex* does).

While our previous discussion has explained how these figures represent a measure of divergence, the numbers themselves mean little without some basis for comparison. Given our ignorance of normal expectations, we might assume that any two passages of hexameter poetry would reveal divergences of equal or greater magnitude, and that our findings have nothing to do with individuality and merely reflect the normal variability of language. The major test of the entire technique was the application of the same test to segments of the same work, and this was done for all points in the line, forward and reverse, for the first and second halves of the *Eclogues* and *Metamorphoses XII.* The contrast in results was truly dramatic. For example, when we compare the first half of *Metamorphoses XII* with the second half, and when we compare *Eclogues 1-5* with *Eclogues 6-10* for words beginning at point 7, the results are :

Met. XII	χ^2 (11 df) = 8.95 (30-40)
Eclogues	χ^2 (9 df) = 8.39 (60-70)

For words ending at point 7 :

Met. XII	χ^2	(7 df)	= 3.12 (10 - 20)
Eclogues	χ^2	(7 df)	= 8.13 (60 - 70)

It is this contrast in results which has convinced me that these findings have inferential as well as descriptive validity. On the basis of the percentiles noted here, no one would ever doubt the null hypothesis that the differences noted are the result of random variation in samples taken from a single population. To strengthen this conclusion, we had also to meet the possible criticism that the comparison of smaller segments will affect the results significantly. Accordingly the same comparison was made between *Eclogues 1-5* and *Met. XII 1-316.* The results for word-types beginning and ending at point 7 are :

χ^2 (10 df) = 29.99 (99.5 - 99.9)
χ^2 (8 df) = 20.79 (99 - 99.5)

It is true that the results are affected by halving the samples, but the dramatic contrast remains. Even so, problems about inference remain.

If we could look upon the distribution in Table 5 as representing two samples of, say, jelly-beans in 14 different colors, then the results arrived at would allow us to say, with as much assurance as the statistical tables allow, that these samples must have come out of two different jars, or parent populations, and that they could not have been taken from the same jar. But what is the nature of the jar or jars with which we are involved in this case ? Further, what we have here are not samples in the ordinary sense of random selections from a well-defined parent population. Given the terrible efficiency of the computer, we have not worked with samples; we have enumerated all the features of the entire available parent population. There is, perhaps, an intuitive sense in which it may be held that these works represent the mode of composition habitual to these authors at the time that they devoted their energies to works of this sort. To the extent that this murky hypothesis may be viewed as not simply tautological, we have something of a confirmatory test in that the above results demonstrate a homogeneity of practice running throughout the entire works. At this stage, it is prudent not to rely too heavily on the paraphermalia of inferential statistics and to take the following stence. First, we may use "if-then" propositions in a legitimately descriptive manner, such as "if these were randomly drawn samples from a well-defined parent population, then we might infer, etc". Second, we can begin to build up a store of empirically derived information from which we may eventually be able to take the inductive leap of inference. Based on the precise circumstances of the test described above, we know that *Metamorphoses XII* when divided in halves will generate a chi-square value of 8.95 with a percentile level between 30 and 40. Just so, the more variable *Eclogues* have a divergence which is measured by a percentile between 60 and 40.

As said above, such slim experience does not comprise much support for inference concerning identity, but we are now a little better off than we were before. Much more experience will make us much better off.

The above discussion has concerned only observations made at point 7. Tables 6 - 8 present the results of our findings at every point in the verse in a comparison of *Eclogues, Metamorphoses XII,* and *Culex.* Tables 9 - 10 present the comparison of first half with second for the *Eclogues* and *Metamorphoses XII.* These results speak for themselves.*

We may conclude our discussion of the hexameter maze with a look at the "raw" chi-square total scores. Since we make no inferential claims for these scores, we must assume our "if-then" posture of description, and we include percentiles on that basis. These scores are repeated here with the first entry in each case being the forward score, the second is the reverse.

Eclogues - Metamorphoses XII	(76 df) 314.62 (99.95 - 100)
	(75 df) 289.41 (44.95 - 100)
Eclogues - Culex	(69 df) 188.04 (99.95 - 100)
	(68 df) 164.84 (99.95 - 100)
Metamorphoses XII - Culex	(64 df) 213.44 (99.95 - 100)
	(66 df) 181.66 (99.95 - 100)
Eclogues l-5 – 6-10	(56 df) 58.96 (60 - 70)
	(55 df) 74.78 (95 - 97.5)
Met. XII 1-314 – 315-628	(50 df) 43.79 (20 - 30)
	(49 df) 37.38 (10 - 20)

* In all cases where an entry is omitted or left blank, there is insufficient information to generate a chi-square value for divergence. This is largely due to the arithmetical limitations of the chi-square test. Either there is no divergence between authors, or the cases of divergence are less than 10 in the two authors combined.

TABLE 6

Eclogues – Metamorphoses XII

	Forward			Reverse		
Point	df	χ^2	Percentile limits	df	χ^2	Percentile limits
0	11	90.27	99.95 - 100			
1	6	20.51	99.5 - 99.9			
1&	1	3.32	90 - 95			
2	3	4.98	80 - 90			
2&	1	0.08	20 - 30			
3	5	33.53	99.95 - 100	7	53.75	99.95 - 100
3&	1	0.80	60 - 70	1	0.00+	2.5 - 5
3+	1	1.45	70 - 80			
4	7	15.90	95 - 97.5	8	31.17	99.95 - 100
4&	1	0.25	30 - 40	1	0.23	30 - 40
5	1	3.53	90 - 95	5	11.92	95 - 97.5
6	1	0.10	20 - 30	2	3.48	80 - 90
6&	1	0.14	20 - 30	1	0.11	30 - 40
7	13	55.46	99.95 - 100	9	39.95	99.95 - 100
8	2	1.78	50 - 60	4	16.21	99.5 - 99.9
9	4	16.04	99.5 - 99.9	2	20.6	60 - 70
9&				1	0.13	20 - 30
10	6	28.43	99.95 - 100	5	1.27	5 - 10
10&	2	1.47	50 - 60	2	4.34	80 - 90
11	2	6.40	95 - 97.5	3	2.35	40 - 50
12	3	12.32	99 - 99.5	9	49.10	99.95 - 100
13	1	0.06	10 - 20			
14	2	17.66	99.95 - 100	4	23.11	99.95 - 100
15	1	0.17	30 - 40	7	47.55	99.95 - 100
I7				4	2.60	30 - 40
TOTAL	76	314.62	99.95 - 100	75	289.41	99.95 - 100

TABLE 7

Eclogue - Culex

	Forward			Reverse		
Point	df	χ^2	Percentile limits	df	χ^2	Percentile limits
0	12	40.61	99.95 - 100			
1	7	4.90	30 - 40			
1&						
2	4	7.72	80 - 90			
2&						
3	4	24.34	99.95 - 100	6	21.41	99.5 - 99.9
3&	2	0.41	10 - 20	3	8.63	95 - 97.5
3+	1	0.11	20 - 30			
4	6	11.79	90 - 95	8	7.78	50 - 60
4&						
5	1	3.33	90 - 95	5	11.42	95 - 97.5
6	2	22.46	99.95 - 100	2	0.69	29 - 30
6&						
7	10	38.62	99.95 - 100	9	34.60	99.95 - 100
8	1	2.79	90 - 95	3	7.20	90 - 95
9	4	7.87	90 - 95	2	0.31	10 - 20
9&				1	0.68	50 - 60
10	6	4.29	30 - 40	5	9.79	90 - 95
10&						
11	2	4.45	80 - 90	2	1.86	60 - 70
12	3	2.20	40 - 50	9	36.01	99.95 - 100
13	1	0.09	20 - 30			
14	2	12.05	99.5 - 99.9	4	4.03	50 - 60
15	1	0.01	5 - 10	6	12.20	90 - 95
17				3	8.22	95 - 97.5
TOTAL	69	188.04	99.95 - 100	68	164.84	99.95 - 100

TABLE 8

Metamorphoses XII - Culex

	Forward				Reverse	
Point	df	χ^2	Percentile limits	df	χ^2	Percentile limits
0	10	37.75	99.95 - 100			
1	5	18.66	99.5 - 99.9			
1&						
2	4	11.58	97.5 - 99			
2&						
3	4	5.28	70 - 80	5	34.85	99.95 - 100
3&	1	1.71	80 - 90	1	0.36	40 - 50
3+						
4	6	16.40	97.5 - 99	7	20.19	99 - 99.5
4&	1	3.05	90 - 95	1	0.46	50 - 60
5				5	3.64	30 - 40
6	2	7.79	97.5 - 99	1	3.81	90 - 95
6&						
7	13	61.34	99.95 - 100	10	20.76	97.5 - 99
8				4	13.29	99 - 99.5
9	4	2.86	40 - 50	1	0.76	60 - 70
9&						
10	6	18.91	99.5 - 99.9	5	12.89	97.5 - 99
10&	1	3.71	90 - 95	1	1.11	70 - 80
11	2	1.95	60 - 70	3	0.57	5 - 10
12	2	16.46	99.95 - 100	9	31.28	99.95 - 100
13						
14	2	0.62	20 - 30	4	17.23	99.5 - 99.9
15	1	0.07	20 - 30	6	14.83	97.5 - 99
17				3	5.62	80 - 90
TOTAL	64	213.44	99.95 - 100	66	181.66	99.95 - 100

TABLE 9

Eclogues 1-5 - Eclogues 6-10

	Forward				Reverse			
Point	df.	χ^2	Percentile limits		df	χ^2	Percentile limits	
0	10	18.49	95 - 97.5					
1	5	5.10	50 - 60					
1&								
2	3	0.33	2.5 - 5					
2&								
3	4	4.41	60 - 70		6	12.13	90 - 95	
3&								
3+	1	0.43	40 - 50					
4	6	4.23	30 - 40		8	12.83	80 - 90	
4&								
5	1	0.08	20 - 30		4	3.86	50 - 60	
6	1	0.55	50 - 60		1	1.16	70 - 80	
6&								
7	7	8.39	60 - 70		7	8.13	60 - 70	
8	1	0.48	50 - 60		3	2.37	40 - 50	
9	4	6.70	80 - 90		1	3.27	90 - 95	
9&								
10	6	4.05	30 - 40		5	6.34	70 - 80	
10&								
11	2	0.56	20 - 30		1	1.73	80 - 90	
12	3	3.91	70 - 80		7	12.26	90 - 95	
13								
14	2	1.24	40 - 50		4	2.09	20 - 30	
15					6	6.08	50 - 60	
17					2	2.53	70 - 80	
TOTAL	56	58.96	60 - 70		55	74.78	95 - 97.5	

TABLE 10

Metamorphoses XII - First and second halves

Point	Forward df	Forward χ^2	Forward Percentile limits	Reverse df	Reverse χ^2	Reverse Percentile limits
0	8	8.70	60 - 70			
1	4	7.29	80 - 90			
1&						
2	3	2.78	50 - 60			
2&						
3	4	3.04	40 - 50	4	5.37	70 - 80
3&						
3+						
4	5	2.90	20 - 30	6	3.43	20 - 30
4&	1	0.96	20 - 30	1	0.23	30 - 40
5				4	8.08	90 - 95
6	1	0.06	10 - 20			
6&						
7	11	8.95	30 - 40	7	3.12	10 - 20
8	1	2.29	80 - 90	3	1.71	30 - 40
9	3	3.26	60 - 70	1	2.39	90 - 95
9&						
10	5	1.78	10 - 20	4	1.29	10 - 20
10&	1	0.17	30 - 40	1	0.16	30 - 40
11	1	0.48	50 - 60	1	1.83	80 - 90
12	1	1.11	70 - 80	8	4.39	10 - 20
13						
14	1	0.004	5 - 10	3	2.59	50 - 60
15				4	1.95	20 - 30
17				2	0.84	30 - 40
TOTAL	50	43.79	20 - 30	49	37.38	10 - 20

These scores make it quite clear that there is less divergence between the *Eclogues* and the *Culex*, than between the *Eclogues* and *Metamorphoses XII*. This is not a surprising result. It is also not surprising, perhaps to find that there is less divergence between *Metamorphoses XII* and the *Culex* than between the *Eclogues* and the Ovidian poem, but the result invites some speculation. Can we conclude, as some critics have, that the *Culex* is an imitation of Vergil which is closer in date to Ovid than the *Eclogues* ? The answer is no. Not yet. Much more experience with this technique is necessary before leaping to any such far-reaching conclusions. The most glaring omission in our experience is that we do not yet know the degree of divergence to be expected between different attested works of the same author. As a first step, it will be most interesting to measure the divergence between the *Eclogues* and a book of the *Aeneid*, but this has not yet been done. It may well be that divergences as great as the ones we have found exist within the corpus of a single poet. The divergence may be due to differences in subject matter or date of composition. We simply do not yet know, and much more remains to be done before we can speak with assurance in this area.

The results we have found thus far are promising. In particular, when we compare the first half of a work with the second half, the resultant "raw" chi-square total score has an accompanying percentile level which is reassuring. The comparative lack of divergence within *Metamorphoses XII* is not surprising. The result for the *Eclogues* is much more interesting. since it is a collection of separate poems. We get, as we should expect, greater divergence than within Ovid's poem, but even in the case of the reverse measurement where the percentile is 95-97.5, the divergence is clearly and measurably less than what we find in comparing the separate works where the percentile level is consistently 99.95-100. All this is suggestive, but no more than that. These results suggest that the "raw" chi-square total is a valid inferentially as the results gained from the tableaus constructed for each point in the line. They also suggest that an appropriate level of significance for statistical work in literature is .001.*

Finally, it must be remarked that in every case but one, the reverse score is lower than its accompanying forward score. This is a very curious fact and invites speculation of the wildest sort. Let us set aside the cases where one work is compared with another, fot it seems fairly obvious that, here this curious result is caused largely by the fact that the reverse measurement must disregard the very large amount of forward variation taking place at point 0. It is much more interesting to think about the comparison of the first half of a work with the second. In the case of Ovid, the reverse score is slightly lower than the forward one in agreement with our general finding. Turning to Vergil, however, we find a sizable disparity in the opposite direction, large enough to raise the percentile level from 60-70 to 95.97.5. The following conclusion seems inescapable : in the *Eclogues*, Vergil is more like himself moving forward through the verse. We may now revert to the question raised long ago in the discussion in which direction does the rat run ?

A good many speculative assumptions are implicit here and it does not really seem worth while to spell them all out, but one of these is no

* See G. Herdan, *Quantitative Linguistics* (London, 1964) 137 F

longer a mere assumption : in writing a poem, authors do evince metrical behavior consistent enough to be termed habitual. We have found, rather to our surprise, that Vergil di plays more consistency if we assume that, in the mysterious process of composition, he is somehow generally moving forward through the verse. The same is not true of Ovid. Here the evidence is weaker, but what there is of it suggests that the compositional process for Ovid was at least as often backwards as it was forwards. The intuitive term "compositional process" is deliberately left vague, and we shall leave it at that.

Our findings are statistical, they are quite precise, and hopefully they have statistical validity. It would be presumptuous, however, to claim that we have anything more than the tiniest hold upon the workings of the mystery of poetic creation. Improper emphasis of findings such as this one is an abomination, but properly understood, we are better off than we were.*

---

* Wild speculation need not cease here. We are not confined to all forward or all backword choices. Depending on the model of composition which we construct (e.g., starting from both ends of the verse and working toward the middle, or the reverse, or innumerable other variants) it is possible to arrive at various formulations which will demonstrate greater degrees of consistency within these authors. We can, for example, find a much greater degree of consistency in Vergil if we assume that he somehow began the compositional process around point 2 and only added the initial word in the verse after the rest of the line was formed. However, the multiplicity of such models (none of which would have more than statistical validity) argues against such approach. The notion of forward as opposed to backward consistency seems general enough to be intuitively useful.

A prelude creates the expectation of a postlude. At the beginning of this discussion we referred to the theme of conscious choice in connection with monosyllables whose length could be controlled, and we came to the conclusion that despite considerable similarities, Ovid had, for example, used these words to emphasize the dactylic quality of his verse. We shall conclude with the description of an aspect of hexameter verse where no individualistic choice whatever is exercised. In using my computer-generated materials, it occurred to me that authors might differ significantly in the degree to which they did or did not use the same words in the same positions in their verse. I resolved to confine my attention only to words which occurred exactly twice. (The forms had to be identical. Thus, *amabat* and *amabant* were classed as different words.)

Since the number of such pairs would depend upon the total length of the text,* I divided my materials into four texts of roughly equal length : *Eclogues* 1-5 (420 lines), *Eclogues* 6-10 (410 lines), *Culex* (414 lines), and *Metamorphoses XII*, 1-414.

I was not surprised to find that the number of pairs occurring in each text was fairly constant :

Text	Lines	Pairs	Pairs/Lines
Met. XII, 1-414	414	216	.522
Eclogues 1-5	420	224	.533
Eclogues 6-10	410	196	.478
Culex	414	217	.524

* See G. Herdan, *Quantitative Linguistics* (London, 1964) 69.

No individuality is exercised here and it seems obvious that we are here dealing with a constant of the language. Even if we assume that the texts are of equal length, the chi-square value is (3df) 2.06 (40-50), and we would not reject the null hypothesis that we have samples from the same population.

It did seem to me quite possible that the texts would differ in the degree to which pairs appeared in the same position in the line. This time my findings revealed an unsuspected constancy :

Text	Total pairs	Same position	Same/Total
Met. XII 1-414	216	83	.384
Eclogues 1-5	224	95	.419
Eclogues 6-10	196	82	.420
Culex	217	93	.410

$$\chi^2 \text{ (3df)} = 0.76 \text{ (10 - 20)}$$

Once again there seems to be no exercise of individuality. I have no ready explanation for these findings. At the moment it seems to me that we are dealing with a feature over which no conscious choice was exercised. In terms of statistical aggregates, whether or not a word recurred in the same position in the line seems to have been of no moment to our authors. What is remarkable is the constancy of the statistical findings.*

* The following acknowledgments are in order. I wish only that they graced a more impressive product. Professor Stephen Waite kinkly supplied me with a machine-readable text of the *Eclogues*. The text of *Metamorphoses XII* was prepared under the auspices of a grant from the Great Lakes Colleges' Association. The text of the *Culex* and the use of a large computer were supplied through the first American Philological Association Summer Institute in Computer Applications to Classical Studies held at the University of Illinois during the summer of 1969. The present study has been written while enjoying the hospitality of Professor Delatte, a stay made possible through the generosity of Oberlin College and the Commission for Educational Exchange between the United States of America, Belgium and Luxembourg.

Oberlin College — Nathan A. GREENBERG