

Virchow, His, Topinard, Broca, Davis, Flower, and a multitude of others, are hopelessly inadequate to maintain the structure reared on them. No valid arguments at all as to race or type can be based on such statistics, and the biometrician can assert this without in the least criticising the appreciative judgment of these great anatomical authorities.

The object of the present series of *Craniological Notes* is to bring home to craniologists the need for the revision of their statistical methods. They will therefore deal principally with recent craniological work, and if they criticise it occasionally with vigour, it will be with the sharpness of the surgeon's knife, which is handled in the real interests of the patient. Our criticism will not be purely negative, but reconstructive.

I. Professor Aurel von Török's attack on the Arithmetical Mean.

By KARL PEARSON, F.R.S.

A very voluminous craniological memoir has recently been issued by Professor Aurel v. Török assisted by Herr Gabriel v. László. It is entitled :

Ueber das gegenseitige Verhalten der kleinsten und grössten Stirnbreite so wie der kleinsten und grössten Hirnschädelbreite bei Variationen der menschlichen Schädelform. "Zeitschrift für Morphologie und Anthropologie," Band iv. S. 500—88.

This memoir consists of two parts, first a vigorous criticism of the use of the arithmetic mean in craniology, and secondly an investigation of the correlation which exists between certain cranial breadths. The data with which the authors deal are very ample, consisting of 2000 skulls "aus älteren und jüngeren Friedhöfen Ungarns." Presumably the crania were adult and of one sex. This first note deals only with Professor v. Török's attack on the use of the arithmetic mean.

Professor v. Török's attack is of the following character. He says that under the expression "type" can only be understood what is characteristic of the totality of any class of things, something which enables us at once to distinguish them from other things:

"So kann...die Bestimmung eines Typus in gar nichts anderem bestehen, als dass man unter dem Mehrerlei der Einzelmerkmale dasjenige herausgreift, was in der grossen Ueberszahl anzutreffen ist. Dies ist doch einfach und klar. Nun muss die Frage gestellt werden: kann durch eine arithmetische Mittelzahl das Charakteristische, d. h. das in der grossen Ueberszahl Vorkommende ausgedrückt werden? Dies wäre nur unter der einzigen Bedingung möglich, wenn auf die Werthgrösse der arithmetischen Mittelzahl die überwiegende Mehrheit der Einzelfälle fallen würde. Entweder, oder. Nun soll man eine solche arithmetische Mittelzahl, welche die überwiegende Mehrheit der Einzelfälle in sich vereinigt, doch endlich einmal auch aufzeigen! Es wurde bisher tausend- und tausendmal der Typus mittels der arithmetischen Mittelzahl schon als bestimmt angegeben, aber es soll noch derjenige Anthropologe ausfindig gemacht werden, der sagen könnte, dass er mit seiner arithmetischen Mittelzahl auch wirklich die Ueberszahl der Einzelfälle zum Ausdruck gebracht hat" (p. 511).

Professor v. Török here clearly considers that the type is something which differentiates an *individual* of one population from an *individual* of a second population. As a matter of fact the arithmetic mean may, but does not always, enable us to distinguish one *race* or population from a second *race* or population. Any such population or race is only *fully* defined by a number of statistical constants, means, modes, variabilities, correlations, etc. and will only be distinguishable from a second race provided its constants one or more or all of them differ by quantities sensibly greater than the probable errors of random sampling from those of the second race. A knowledge of the arithmetic means only would never enable one to say of an individual skull that it belonged to one race and *not* to another. A knowledge of the *variabilities and correlations as well* of these races might enable us to state *the degree of probability* that the skull belonged to one race rather than to the other.

If it is necessary for Professor v. Török at this epoch in the history of craniology to define the "type" as that which is the rule, and to demonstrate against the great body of craniologists that the arithmetic mean does not define the type, this science must indeed be in a parlous condition! One can hardly grasp that this can be the state of affairs in any branch of knowledge, or that its workers have never heard of Quetelet, Gulton, or even Stieda!

But Professor v. Török perceives that he cannot leave the matter here. He states that the arithmetical mean according to the doctrine of Gauss is that value of the character which most frequently repeats itself in the population, i.e. it coincides with what biometricians are accustomed to term the *mode*. Our authors, merely taking the rough frequencies of their 2000 skulls, show that in three cases out of four the group which corresponds to the arithmetic mean falls short of other groups of frequency in its neighbourhood, and accordingly argue that the arithmetic mean is idle:

Wie wir sehen, trifft der Fall der häufigsten Wiederholung (Vertretung) der Werthgrösse der arithmetischen Mittelzahl bei den vier Variationsreihen nur ein einziges Mal ein. In der Variationsreihe der grössten Hirnschädelbreite wird sie sogar durch vier andere Einzelwerthe übertroffen. Man frage sich doch angesichts dieser Thatsache, was man Sicheres auf den Beweis einer arithmetischen Mittelzahl behufs unserer Typusforschungen bauen kann? Und wenn man ohne Voreingenommenheit die Sache so wie sie ist beurtheilt, muss man denn nicht zur Einsicht gelangen; dass eben, weil ihre Werthgrösse nie die absolute Mehrheit und auch nicht immer die relative grösste Vertretung aufweist, wir genöthigt sind behufs Aufstellung eines Typus ausser ihr unbedingt noch andere Werthgrössen des betreffenden Masses in Betracht zu ziehen! (S. 517).

Of course other constants must be dealt with in addition to completely describe a given population. Only Professor v. Török's attack fails absolutely, because he has neglected to deal with the variations in his frequency groups due to random sampling, and these as we will show later are such that there is no really *significant* difference between the actual modes,—not his apparent modes (*Scheingipfeln*),—and his means.

Lastly, Professor v. Török says that the position of the arithmetical mean in the whole group of variation must be considered.

Wie schon der Ausdruck "Mittelzahl" andeutet, erweckt der Begriff einer solcher Zahl in uns die Vorstellung, dass ihre Werthgrösse in der Mitte der betreffenden Variationsreihe (Zahlreihe) steht. Thut sie aber dies? Wenn dies der Fall wäre, so böte sie uns innerhalb der so vielen Veränderlichkeiten unserer kranimetrischen Zahlreihen wenigstens einen fixen Punkt, wo man das Problem einer Typusbestimmung anheben könnte (S. 517).

In other words, to use the biometric term, Professor v. Török questions whether the mean gives the median. By the simple process of taking the characters of his first and last groups of variants adding and halving them, Professor v. Török shows that the arithmetic mean does not give the median! There is not the least attempt to obtain a scientifically accurate median, nor to ascertain whether the difference between that and the mean is a sensible difference or not. The writer has clearly not the most elementary conception of the theory of statistics, nor even a merely physical appreciation of the effects of random sampling, nor again of differences due to different methods of grouping. The whole problem of graduating raw data is a sealed book to him! Yet here are his final conclusions,—sweeping for craniology, revolutionary, indeed, for all science!

Auch nach den hier vorgetragenen Thatsachen, muss ein jeder selbständig denkender Forscher die Verfehltheit aus der arithmetischen Mittelzahl, einen Typus herausspeculiren zu wollen, doch einsehen.... Es erwächst nun eine unabweisliche Pflicht, und zwar in erster Reihe für die tonangebenden Autoritäten der heutigen Anthropologie: entweder, wenn möglich, ihr bisheriges Verfahren klar und einwurfsfrei zu rechtfertigen; oder, wenn dies ihnen nicht gelingt, mit gutem Beispiel voranzugehen und die Typusfrage von diesem alten Irrthume endlich einmal zu befreien. Ein weiteres Still-schweigen ist gewiss nicht mehr motivirt, die Angelegenheit muss doch einmal geschlichtet werden. Der Kampf um die Wahrheit ist das einzig berechtigte *punctum saliens* für wissenschaftliche Discussionen, neben welchem alle übrigen persönlichen Rücksichten verstummen müssen! (S. 520).

Let us now see exactly in what Professor v. Török's criticism consists. It contains three separate charges against the arithmetic mean :

(i) It does not give the rule, something characteristic of every member of a group termed by craniologists the type.

Obviously and clearly it cannot, and if any craniologist thinks it does, he should begin his studies *ab initio* with a reading of Quetelet's works.

(ii) It is not identical with the *mode*.

(iii) It is not identical with the *median*.

These defects (i), (ii), and (iii) make the mean of no service at all in craniological discussions.

Now let us consider (ii) and (iii) in the light of Professor v. Török's own data treated by an adequate statistical theory. Table I. gives his data for greatest forehead breadth (S. 509) and for greatest skull breadth (S. 509). I selected these two series out of the four given because they looked in Professor v. Török's diagrams the "skewest," and therefore, if there was a sensible distinction between mean, mode and median, I thought it would certainly be evidenced in these.

TABLE I.

GREATEST FOREHEAD BREADTH				GREATEST SKULL BREADTH			
mm.	Frequency	mm.	Frequency	mm.	Frequency	mm.	Frequency
100	1	123	122	120	1	145	128
101	0	124	109	121	0	146	118
102	0	125	80	122	0	147	115
103	4	126	81	123	0	148	84
104	5	127	60	124	1	149	78
105	5	128	45	125	0	150	52
106	3	129	46	126	0	151	46
107	14	130	33	127	6	152	29
108	13	131	26	128	1	153	30
109	21	132	17	129	9	154	23
110	33	133	19	130	13	155	19
111	50	134	4	131	20	156	8
112	43	135	7	132	28	157	5
113	49	136	1	133	36	158	3
114	89	137	2	134	36	159	5
115	99	138	0	135	58	160	4
116	108	139	1	136	60	161	0
117	124	140	0	137	88	162	1
118	114	141	0	138	108	163	1
119	151	142	0	139	118	164	0
120	148	143	1	140	119	165	0
121	141	144	0	141	143	166	0
122	130	145	1	142	138	167	0
				143	127	168	0
				144	140	169	1

Now these series are so nearly symmetrical that it seemed sufficient to graduate them with a curve of the type : $y = y_0 \left(1 + \frac{x}{a}\right)^p e^{-rx}$, where $p = \gamma a$: see *Phil. Trans. A.*, Vol. 186, p. 387.

We find :

<i>Greatest Forehead Breadth.</i>	<i>Greatest Skull Breadth.</i>
Mean = 120.0475 mm.	Mean = 142.7095 mm.
S.D. = 5.8124 mm.	S.D. = 5.8374
μ_1 = 33.783,910	μ_1 = 34.075,776
μ_2 = 11.370,311	μ_2 = 17.810,678
p = 1192.0125	p = 497.9270
y = 5.9425	y = 3.8264
a = 200.5918	a = 130.1279
y_0 = 137.273	y_0 = 136.684
Mode = 119.8792 mm.	Mode = 142.4482 mm.

The mode being the origin the equations to the frequency distributions are respectively :

$$y = 137.273 \left(1 + \frac{x}{200.5918} \right)^{1192.0125} e^{-5.9425x},$$

and :

$$y = 136.684 \left(1 + \frac{x}{130.1279} \right)^{497.9270} e^{-3.8264x}.$$

The high values of p and a show how very close the distributions are to the Gaussian curve in which mean, mode, and median coincide. Let us consider this matter a little more at length by finding the probable errors of μ_2 and d , the distance between mean and mode. The values of these probable errors are to be found in the paper on the *Probable Errors of Frequency Constants**, and we thus find for :

<i>Greatest Skull Breadth.</i>	<i>Greatest Forehead Breadth.</i>
$\mu_2 = 11.3703 \pm 17.3500$	$\mu_2 = 17.8107 \pm 7.5783$
$d = .1683 \pm .1081$	$d = .2163 \pm .1088$

Thus μ_2 is in one case slightly more than double the probable error and may, perhaps, be significant. In both cases d is less than twice its probable error and is therefore very likely to be insignificant. Actually this is the slender sort of basis upon which Professor von Török would discard the arithmetic mean from craniology! Looking at the result a little more closely, we see that Professor v. Török having no conception of the probable error of any sub-group of frequency places his mode at the *Scheingipfel* of 121 mm. for the forehead breadth. The actual mode of the graduated observations is at 119.88 mm., only .17 mm. from the mean at 120.05 mm.; and the difference between the two is really within the limits of the errors of random sampling. Turning to the maximum skull breadth, the actual mode is at 142.45 mm. and the mean at 142.71 mm., while Professor v. Török points with triumph to *Scheingipfel* at 141 and 144 as if they subverted the whole theory of the arithmetic mean! Again we see that the true mode as obtained from a proper system of graduation is quite close to the mean, i.e. within a small fraction of a millimetre, and the difference is quite possibly due to the errors of random sampling.

In order to illustrate these points still further the normal curves for the two series were calculated; they are :

<i>Greatest Forehead Breadth.</i>	<i>Greatest Skull Breadth.</i>
$y = 137.273 e^{-\frac{1}{2}\left(\frac{x}{5.8124}\right)^2}$	$y = 136.684 e^{-\frac{1}{2}\left(\frac{x}{5.8374}\right)^2}$
Origin at mean: 120.0475 mm.	Origin at mean: 142.7095 mm.

* *Biometrika*, Vol. II. p. 273. See also *Phil. Trans. A.*, Vol. 191, p. 275.

Table II. gives the observed and calculated values. Applying the method for testing goodness of fit given in *Biometrika*, Vol. I. p. 155, we find in the first case: $\chi^2=40.429$ and $P=.37$, and in the second case, $\chi^2=33.362$ and $P=.69$. Or, supposing the Hungarian data dealt with by Professor v. Török to actually follow the normal or Gaussian curve, in every three samples of 2000 skulls he would on the average have found one fitting theory worse for forehead breadth than his sample actually does; and for skull breadth every two out of three samples of 2000 skulls would on the average give a worse result. In the face of such probabilities as this any sound statistician would not hesitate to say that for skull breadths Professor v. Török's Hungarian data obey the normal law of frequency, exactly as we have proved is the case for many series of other cranial measurements. The accompanying *diagrams* show the observations fitted with the theoretical curves and demonstrate at a glance how idle is any argument based on *Scheingipfeln*.

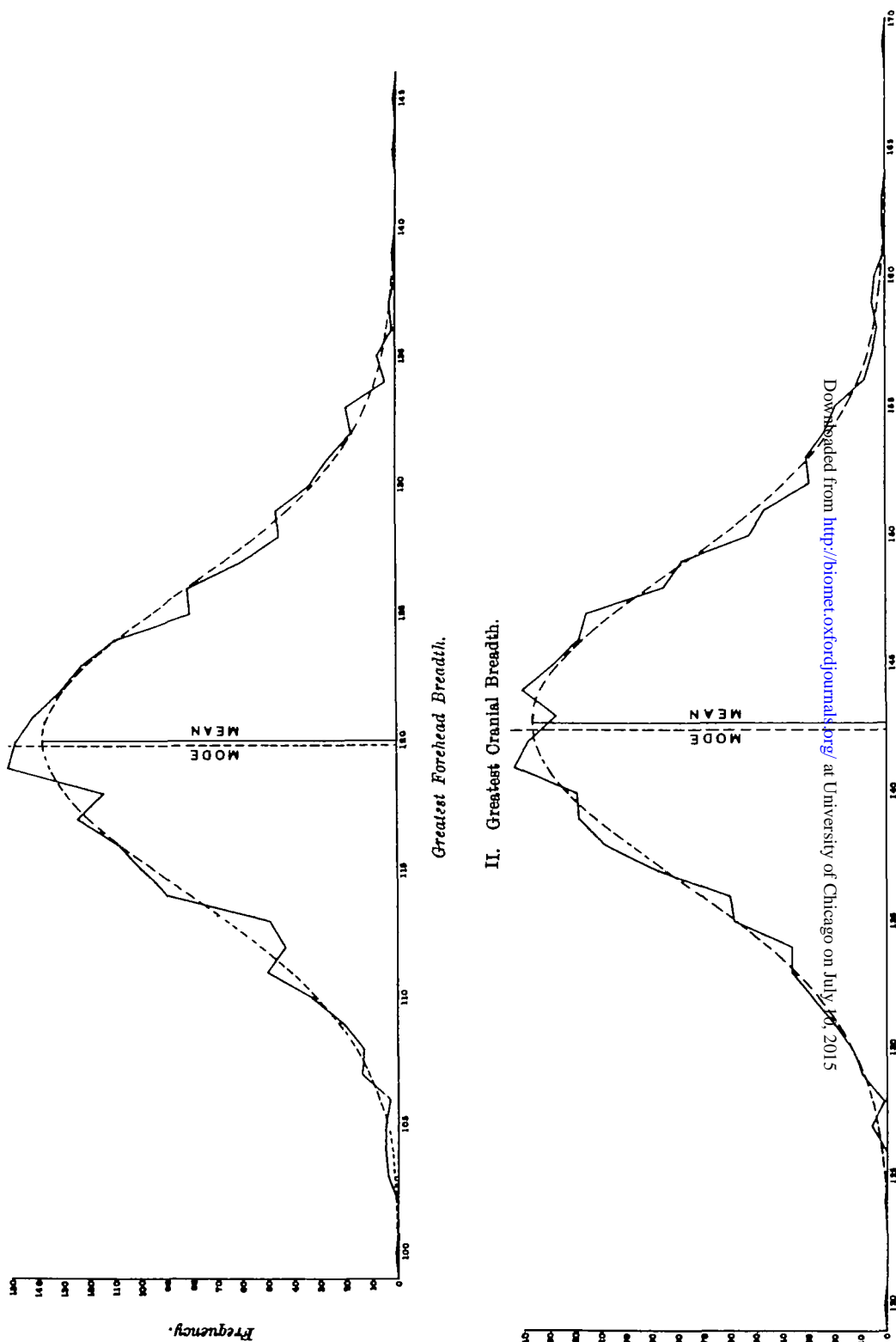
TABLE II.

mm.	GREATEST FOREHEAD BREADTH					mm.	GREATEST SKULL BREADTH				
	Frequency		mm.	Frequency			Frequency		mm.	Frequency	
	Observed	Calculated		Observed	Calculated		Observed	Calculated		Observed	Calculated
Below 103	1	2.5	125	80	95.5	Below 127	2	5.5	148	34	90.6
103	4	1.9	126	81	81.2	127	6	3.7	149	78	76.5
104	5	3.1	127	60	67.2	128	1	5.8	150	52	62.7
105	5	4.8	128	45	53.9	129	9	8.7	151	46	49.9
106	3	7.4	129	46	42.0	130	13	12.9	152	29	38.6
107	14	11.1	130	33	31.8	131	20	18.3	153	30	29.0
108	13	16.0	131	26	23.3	132	28	25.5	154	23	21.1
109	21	22.7	132	17	16.6	133	36	34.3	155	19	15.0
110	33	30.9	133	19	11.5	134	36	45.0	156	8	10.3
111	50	40.9	134	4	7.8	135	58	57.2	157	5	6.6
112	43	52.7	135	7	5.1	136	60	70.6	158	3	4.7
113	49	65.8	136	1	3.2	137	88	84.7	159	5	2.8
114	89	79.9	137	2	2.0	138	108	98.7	160	4	1.7
115	99	94.2	138	0	1.2	139	118	111.6	161	0	1.1
116	108	107.7	139	1	0.7	140	119	122.6	162	0	0.6
117	124	119.5	above			141	143	130.8	163	1	0.3
118	114	128.9	139	2	0.8	142	138	135.5	above		
119	151	134.9				143	127	136.0	163		0.4
120	148	137.1				144	140	133.5			
121	141	135.2				145	128	126.4			
122	130	129.6				146	118	116.5			
123	122	120.6				147	115	104.3			
124	109	108.8									

Had the mean differed from the mode, modern statistical theory could have deduced the modal value without appealing to *Scheingipfeln*. As it turns out, however, for these skulls' breadths Gaussian theory is perfectly applicable: *mean, mode and median sensibly coincide*, and Professor v. Török's attack on the arithmetical mean fails absolutely, and this because he is attempting, as so many other craniologists do, to advance without any knowledge of statistical theory.

If the median were to coincide with the mean Professor v. Török tells us that there would be a fixed point to start the "Problem einer Typusbestimmung" from. Well, it actually does

2000 ♂ Hungarian Skulls. I. Greatest Forehead Breadth.



coincide in Professor v. Török's data within the errors of random samplings. The arithmetic mean can be freed from the charges he brings against it. Can he solve the "Problem einer Typusbestimmung" on these lines? Personally I think not. "Der Kampf um die Wahrheit" in the craniology has for its solely legitimate *punctum saliens* the recognition that craniology is a branch of biometry, and can only be followed profitably when the modern theory of statistics has been properly studied.

II. Homogeneity and Heterogeneity in Collections of Crania.

By KARL PEARSON, F.R.S.

In a review* by Mr C. S. Myers of a memoir by Miss C. D. Fawcett†, one of the biometric workers at University College, London, exception is taken to the arguments adduced in favour of the homogeneity of the Naqada prehistoric crania. There are many statements in Mr Myers' review which it would be easy to traverse, but as the writer expresses himself as in sympathy with what he is pleased to term the "new path," a consideration of one of them will suffice to show to what extent he appears to be fit at present to enter upon it.

The statement is as follows :

The question, moreover, arises, are we entitled to consider either the Naqada or the above English skulls as belonging to people of a single race? The authors think that we are "justified in treating our material as homogeneous and in speaking of a Naqada *race* and not merely of the Naqada crania" (p. 424). "If the [Naqada] material were markedly heterogeneous the variability in length and breadth of skull ought to be large as compared with admittedly homogeneous material" (p. 424). The standard deviations of the male Naqada skulls and of Bavarian, Aino, French, and English male series as regards skull-length are 5.722, 6.088, 5.936, 7.202, 6.446 respectively, and as regards skull-breadth are 4.612, 5.849, 3.897, 6.068, 4.976 respectively. These deviations (and the "coefficients of variability" derived therefrom) are considered by the authors to be small enough to warrant the conclusion that the Naqada crania, the old Bavarian crania of Professor Ranke, and the Whitechapel English crania of Professor Thane constitute each a homogeneous series. They have left neglected the question whether a much larger standard deviation would result, were we to consider a series, say, of forty-nine male skulls of most diverse ethnic types, composed, *e.g.*, of fifteen Australians, seven Guanches, fifteen Eskimos, and twelve Chinese. This is the material which the reviewer set himself to work out, taking the data haphazard from Flower's well-known catalogue of skulls in the Royal College of Surgeons' Museum. The results gave him a standard deviation of 8.389 for the skull-length, and of 7.002 for the skull-breadth. We see, then, how small is the difference of variation between the Naqada skulls of Professor Pearson's series (which are of a "homogeneous character") and a series which is as heterogeneous as it could well be. Are we, then, not justified in considering the Naqada skulls and the others of Professor Pearson's series as if they had sprung from a mixture of races? If not, at least the problem is less simple than the writers appear to think.

The mean of the variabilities of the skull lengths given in the above paragraph is 6.2788 and of the skull breadths is 5.0804. Mr Myers mixing Australians, Guanches, Eskimos and Chinese finds a variability of skull length = 8.389 and of skull breadth = 7.002. He then points to the differences (2.1102 and 1.9216) and triumphantly asks how such small differences can be of any importance! But had Mr Myers had a mathematical training he would know that nothing is "small" absolutely, but only relatively to something else, and had he had a statistical training he would have known that he must compare it with the variability of these variabilities *i.e.* the standard deviation of the standard deviations of the skull measurements. Now the standard deviation of the above series of skull length variabilities = 5.185 and that of the

* *Man*, February, 1903, p. 13.

† *Biometrika*, Vol. i. pp. 408-467.