On the homogeneity of
the European Eel population (*Anguilla anguilla*)

E.F. Harding
Statistical Laboratory, Cambridge University, 16 Mill Lane, Cambridge CB2 1SB, U.K.

Abstract

The paper considers evidence bearing on the homogeneity of the European *Anguilla anguilla* population in respect of total numbers of vertebrae (TNV). It is in two parts. In Part 1, data and conclusions published by Johs. Schmidt in 1912 and 1913 are examined, and evidence of unusual homogeneity is found. These data are then compared (a) with data known to Schmidt but not published by him, and (b) data from specimens left by Schmidt but not worked up by him. Significant inhomogeneity is found. Interpretation is ambiguous, but it is concluded that Schmidt's claim (1913) of complete homogeneity and a unique spawning place appears as an over-statement. In Part 2, data on length and TNV for some very large samples from Højer in Denmark are analysed. It is found that the material can be represented as a mixture of at least two, possibly three, distinct groups, each with its own distribution of length and TNV. As a final conclusion, the European eel population is not homogeneous as claimed by Schmidt, and Schmidt had in hand at least some of the evidence for this. There are several possible explanations (amongst which, multiple spawning grounds).

Contents

Part 1

Materials

Published sources are the data on total number of vertebrae (TNV) for *Anguilla anguilla* in Schmidt (1913) and in Boëtius (1980). Data from these sources is not repeated in detail; refer to the original publications. The notation PTn denotes

Sample No. n in the Primary Table of Boëtius 1980: thus PT43 refers to a sample of 133 specimens from Ravenna in Italy (published by Schmidt (1913) as col. 11 of table IV). The notes on pp. 108-110 of Boëtius (1980) should also be consulted. Published counts of total numbers of vertebrae have all been made using Schmidt's principle for resolving difficulties with caudal structures (Schmidt 1913, p. 7; Ege 1939, p. 8; Boëtius 1976, p. 206, and 1980, p. 95).

Specimens of '*A. anguilla*' whose TNV count was less than 110 have not been taken into consideration.

Reference is also made to an unpublished laboratory notebook of A. Strubberg, dated 1911, which is clearly the primary document in which are recorded details of samples (dating from 1906-1912) used by Schmidt for his 1912 and 1913 publications. This notebook is in the possession of Dr J. Boëtius, and I am grateful to him for access to it and for permission to refer to it in this paper.

Preliminary considerations

In 1912 Johs. Schmidt compared vertebra counts (TNV) for samples of eels (a) from places on the Mediterranean seaboard, (b) from places on the Atlantic. He stated that the average TNV in ca. 1000 Mediterranean eels was 114.736, and the average in ca. 1700 Atlantic eels was 114.731, and continued: 'A more complete agreement can hardly be imagined and we see, that there is absolutely nothing in the way of our conclusion, that the Mediterranean eels come from the Atlantic. At the same time, these results give the first evidence that only one species of *Anguilla* occurs in North and South Europe and on the islands of the Eastern Atlantic, a fact previously accepted but by no means proved.' (Schmidt 1912, p. 324).

In Schmidt (1913, table IV) are found details of 5 Mediterranean samples (980 specimens, mean TNV = 114.736) and 11 Atlantic samples (1795 specimens, mean TNV = 114.724), closely agreeing with the above. It seems possible that the data used for the 1912 statement is the same as given in 1913, except for one sample taken from the Azores in 1912 (given by Boëtius, 1980, Primary Table sample no. 35: 101 specimens; leaving 1694 Atlantic specimens with mean TNV = 114.730).

Such agreement, more like Physics than Biology, invites scrutiny. Standard deviations for the various samples in Schmidt (1913) vary slightly around $\sigma = 1.3$, so that the S.D. for the difference of the means is 0.052; thus the difference of means is 0.096 standard deviations in 1912, and 0.23 standard deviations in 1913. The chance that so *small* a difference could arise for two samples even from identical populations is $P = 0.18$ for 1913, and $P = 0.076$ for 1912. Though not quite 'significant' even for 1912, the close agreement is interesting.

Pursuing the same point, an analysis of variance on the 16 samples from Schmidt (1913) is done next.

For the 11 Atlantic samples from Schmidt (1913) table 4, we have

	S.S.	d.f.	Mean S.S.
Between sample means	8.231	10	0.822
Within samples (resid.)	3010.7	1784	1.688

$$F = 0.487; \quad P(F_{10:1784} \leqq 0.487) = 0.104$$

Thus, even supposing that eels of the Atlantic area are truly homogeneous, the samples published by Schmidt still form an unusually homogeneous set.

For the 5 Mediterranean samples from Schmidt (1913) table 4, we have

	S.S.	d.f.	Mean S.S.
Between sample means	10.043	4	2.511
Within samples (resid.)	1646.5	975	1.689

$$F = 1.487; \quad P(F_{4;975} \leqq 1.487) = 0.797$$

which is within the acceptable middle range of the F distribution.

The published Mediterranean material therefore shows no evidence of undue homogeneity – nor of *in*homogeneity.

The situation changes when material not published by Schmidt is considered. This is done in the following section.

Material not published by Schmidt

The source for this is Boëtius (1980), Primary Table. We study material in his Categories II (material worked up by Schmidt but never published, as ascertained from protocols etc.) and III (material obtained by working up, for the first time, the collection of preserved specimens left by Schmidt).

Consider first the Mediterranean material: 5 samples in Category II (PT42, PT46, PT48, PT49 and PT50) and one (PT47) in Category III, all dating from 1920 and later. It will be analysed as one lot, since the one III-sample does not seem discordant with the others. As above, we have

	S.S.	d.f.	Mean S.S.
Between sample means	24.709	5	4.942
Within samples (resid.)	2301.4	1405	1.638

$$F = 3.017; \quad P(F_{5;1405} \leqq 3.017) = 0.010$$

Therefore the 6 unpublished Mediterranean samples (5 of them available to Schmidt) are significantly inhomogeneous at the 1% level of significance. Thus although the material (as published) available in 1913 was consistent with a conclusion that the Mediterranean population was homogeneous, if Schmidt had re-examined the question after 1920 using his own data he might have formed a different conclusion.

It must also be said that, from the Strubberg Notebook (which deals with samples published by Schmidt, 1913), the raw counts were subject to repeated or comparative counting (by Schmidt and Strubberg), and there was evidently some debate about whether certain doubtful specimens should be included (fused vertebrae etc.), and about what to do when comparative counts conflicted. Possibly the unpublished data were not so treated, and are less homogeneous in consequence. It is impossible to tell whether the refinement applied to published data has imposed excess uniformity, or whether unrefined unpublished counts are excessively variable.

We next consider unpublished material from the Atlantic, in the Schmidt collections. A fair amount of this consists of samples taken not later than 1913, and it is appropriate to consider this separately. It is highly heterogeneous. Table 1 shows

Table 1. Samples of Atlantic material not published by Schmidt, up to 1913 (identified by Primary Table number from Boëtius, 1980), differences from overall mean of mean vertebra count as multiples of the S.D. of the difference, sample category (II or III, see text), sample sizes, and geographical positions.

No.	Deviation		Category	Number of specimens	Provenance
PT1 :	−8.62	S.D.	III	48	Iceland
PT2 :	−1.17	−	II	15	Iceland
PT4 :	−0.40	−	III	162	Iceland
PT9 :	−1.27	−	III	24	Faroes
PT11:	−10.30	−	III	144	Faroes
PT12:	−8.60	−	III	349	Faroes
PT13:	−0.43	−	III	41	Faroes
PT14:	−0.36	−	III	83	Faroes
PT17:	−0.06	−	II	31	Orkney Is.
PT18:	−3.56	−	III	58	Orkney Is.
PT19:	−6.88	−	III	42	Hebrides
PT20:	−5.28	−	III	13	Norway
PT39:	+3.55	−	II	2	Canary Is.
PT40:	+0.95	−	II	22	Spain

the deviation of the mean of each sample from the mean of the Atlantic material published by Schmidt, as a multiple of the calculated standard deviation for the difference of means in each case.

It is quite unnecessary to make a formal test of differences of the means. Most (but not all) of the most deviant means are III-samples; many (but not all) III-samples are deviant.

Heterogeneity and 'races'

It is established above that consideration of data not published by Schmidt implies heterogeneity. The problem now is to interpret this in terms of possible 'local races', and we face the difficulty of determining what degree of heterogeneity might be required for this. From one point of view, even slight heterogeneity, once established as above, could be taken as evidence against the hypothesis that the European eel population has been indiscriminately distributed over the seaboard, from a common compact breeding ground. On the other hand, some slight degree of variation could arise through differences of environment encountered during earlier development along different routes to the coasts, compatible possibly with dispersal from a more extended, but still unified, breeding area.

The question is therefore one of somehow 'calibrating' the heterogeneity in terms of possible distinct 'races'. This cannot be done from the *A. anguilla* data alone, without arguing in a circle, unless there is a clear geographical component in the variation of sample mean values (but a possible genetic origin is still an open question).

As to a geographical component, it is clear from the above tabulation of Atlantic material that 'Northern' samples usually are markedly deviant, negatively, from

Table 2. Samples of Mediterranean material identified by Primary Table number from Boëtius (1980), Categories I (published by Schmidt), II and III (see text), mean TNV, sample dates and geographical positions.

Sample	Category	Mean TNV	Date	Geographical position
PT47:	III	114.535	1931.05.16	44°41′N 12°10′E NE Italy
PT42:	II	114.624	1930.01.25	36°51′N 10°17′E Tunisia
PT44:	I	114.626	1911.01.27	43°32′N 10°18′E NW Italy
PT51:	I	114.753	1911.05.13	35° – ′N 33° – ′E Cyprus
PT45:	I	114.772	1911.02.23	44°41′N 12°10′E NE Italy
PT50:	II	114.784	1922.01.01	31°08′N 29°51′E Egypt
PT43:	I	114.835	1906. – . –	44°28′N 12°15′E NE Italy
PT49:	II	114.835	1920.12.18	31°08′N 29°51′E Egypt
PT46:	II	114.885	1922.12. –	43°22′N 10°18′E NE Italy
PT41:	I	114.888	1911.01.26	43°25′N 03°42′E France
PT48:	II	114.917	1920.02. –	31°16′N 32°18′E Egypt

the general mean. Boëtius (1980) has established a clear if small gradation from 'Northern' (Iceland, Faroes, Orkneys, Hebrides, Scandinavia: mean TNV = 114.465) through 'Central' (England, Ireland, Atlantic France and Spain: mean TNV = 114.512) to 'Southern' (Azores, Madeira, Canaries, Mediterranean lands: TNV = 114.746) (S.D. = ca. 0.02 in each case; specimens with TNV $\leq$ 110 omitted). Further, only in the 'Northern' region do sporadic specimens with TNV = 105-110 occur. He interprets these results as partly due to possible mixing of *A. anguilla* and *A. rostrata* in the southern part of the 'spawning area', to possible hybridisation, and to possible variation in the arresting of vertebral development.

The means of the 11 Mediterranean samples are presented in Table 2, in increasing order, with the sample dates and geographical positions. (Category I denotes samples published in Schmidt (1913).)

There is no systematic association between mean TNV and date, geographical position or category. Therefore the degree of variation between sample means in the Mediterranean material, established as significant above, may be compatible with non-existence of local races.

Comparison with other species – Zoarces viviparus

Schmidt himself has carried out studies of racial variation, and its genetic and environmental components, especially in the species *Zoarces viviparus* (eel-pout or vivparous blennie) (Schmidt 1917-1930). This species has a range of TNV embracing that of the Atlantic eel species *A. anguilla* and *A. rostrata* jointly, and that respect is an apt comparison. There is no doubt that it forms local breeding groups, and the TNV differences between these reflect true local races. This work, in its bearing on the eel, has been carefully discussed by Harden Jones (1968, pp. 81-83).

For various reasons, however, it is here preferred not to pursue the comparison with *Z. viviparus*. *Zoarces* has many features favoring racial investigations, being sedentary – so mating will be only locally random; live-bearing – so no eggs or larvae will drift off to other localities; and easily obtained at all stages of the breeding cycle so that experiments of all kinds can be carried out. The eel is quite different in all these respects.

Nonetheless Schmidt has drawn definite conclusions about the eel from the results for *Zoarces*. 'A main result of our variational-statistical comparison of the *Zoarces* samples with those of the eel will thus be the following: That *Zoarces vivi-parus* in the north of Europe is divided up into numerous distinctly different stocks of populations according to locality, whereas all the eels of Europe are identical.' (Schmidt 1915, p. 21). 'With regard to the cause of this homogeneity in all samples of *Anguilla vulgaris* and of the heterogeneity in the *Zoarces* samples, there is little room for doubt. The difference must be due to the fact that all European eels are of the same origin, whereas the *Zoarces* exist as separate communities each distinct from the other.' (Schmidt 1915, p. 22).

Whatever the case for *Zoarces*, the logic of this conclusion for the eel is curious. If systematic and stable variation for the eel, comparable to that of *Zoarces*, had been observed, it would of course have reduced the probability of common origin; it was not. However, the eel is highly mobile as larva, elver and migrating adult; and a more or less promiscuous mingling could occur, of individuals spawned in different places.

Conclusions for Part 1

It is established that Schmidt's published material on vertebra counts (TNV) is extremely homogeneous.

Unpublished material is however significantly non-homogeneous. In the case of material relating to the Atlantic seaboard, mean TNV increases from North to South, and the trend continues into the Mediterranean Sea.

Within the Mediterranean Sea, the material is again significantly non-homo-geneous, but there is no apparent correlation with geographical position. Such variation could therefore be consistent with non-existence of local 'races', or poss-ibly with intermingling of sub-populations arriving at the same place from dif-ferent spawning-grounds or by different routes.

The differences between Schmidt's published material (Category I) and material of Category II (unpublished, but known to Schmidt) suggests that either the pub-lished material became more uniform than it should have been in the course of preparation for publication, or the unpublished material was not subjected to the same care and scrutiny as the published.

Category III material (unknown to Schmidt, but worked up for the express purpose of publication in Boëtius (1980)) differs from both Categories I and II. This could be due to differences between workers. On the other hand, every effort was made to count 'according to Schmidt'. Possibly, then, (a) the differences between I and III are genuine, and reflect true inhomogeneity, while differences between I and II, and II and III, reflect only that II was not worked up with normal care for publication. Alternatively, Category I could be aberrant. Finally, it must be borne in mind that Category III material had been preserved for several decades before examination.

The evidence is, therefore, ambiguous. Taken all together, however, it does indicate that Schmidt's claim of homogeneity, and that there is but one spawning place, appears as an overstatement.

Part 2

Materials

In this part we study data published by Boëtius (1976). Here we have by far the largest known set of *Anguilla anguilla* material from one place for which vertebra counts have been made (see Boëtius 1980).

Boëtius (1976, Primary Table III) presents vertebra counts (TNV) and lengths for 4310 elvers in three samples (I, II and III) taken at Højer in Jutland in 1972 (out of 22 934 specimens taken). Samples I, II and III were taken on 1972, April 24, May 14 and June 8 respectively, i.e. at the beginning, middle and end of the elver run. Each sample was subdivided into sub-samples according to apparent developmental stages A-E based on pigmentation according to criteria developed by Gilson (1908) and Strubberg (1913), and the following were chosen for TNV counting: IA, ID, IIA, IID, IIE, IIIC and IIIE. The cited Primary Table should be corrected in that one specimen (ID, 65 mm, TNV = 113) should be removed and one specimen (IIIE, 55 mm, TNV = 114) should be added (J. Boëtius, private communication).

In making the TNV counts, the principle of Schmidt (1906, p. 240; 1913, p. 7) was used to give a definite count, avoiding uncertainty due to variability of the caudal structures: 'the short atlas was counted as no. 1 and the last hour-glass shaped vertebra was taken as the next but last vertebra' (Boëtius 1976, p. 206). This principle is illustrated diagrammatically by Schmidt (1913, fig. 1).

Table 4 of Boëtius (1976) gives the distribution of TNV for the above-mentioned sub-samples. For the present purpose we ignore the 20 specimens with TNV $\leqq$ 110, since most or all are compatible with criteria for attribution to the species *Anguilla rostrata*. On the other hand few if any specimens with TNV = 111 should be so attributed. We ignore also the single specimen (noted above) with length = 55 mm, all others being at least 61 mm. Therefore we deal with data on TNV and length for 4289 specimens of elvers from Højer. These specimens are entirely '0-group' elvers, i.e. have metamorphosed within the preceding year.

In the same publication, Boëtius presents data for 2150 specimens from Esrom in Sealand. Beyond making some use of the relationship between length and TNV noted in that publication, we do not analyse this sample in detail here, since it is a mixed sample of elvers ('0-group') and small yellow eels ($\geqq$ 'I-group'). Data for the '0-group' alone cannot be identified in the Primary Table.

Lengths of elvers are total lengths to 1mm below.

TNV-Length relationship

Boëtius (1976, fig. 6) shows for the Højer samples that mean TNV varies with length (L), increasing from 113.3 when $L = 62$-63 to 115.1 when $L = 78$-79. In Fig. 1, A-F we present the converse relationship, giving mean length for each TNV for sub-samples IA, ID, IIA, IID, IIIC and IIIE (IIE is a very small group). Again it is clear that in each case the relationship is increasing, and approximately linear. The slopes are similar across sub-samples: that is, the relationship is broadly independent of the phase of the elver run, and of developmental stage.
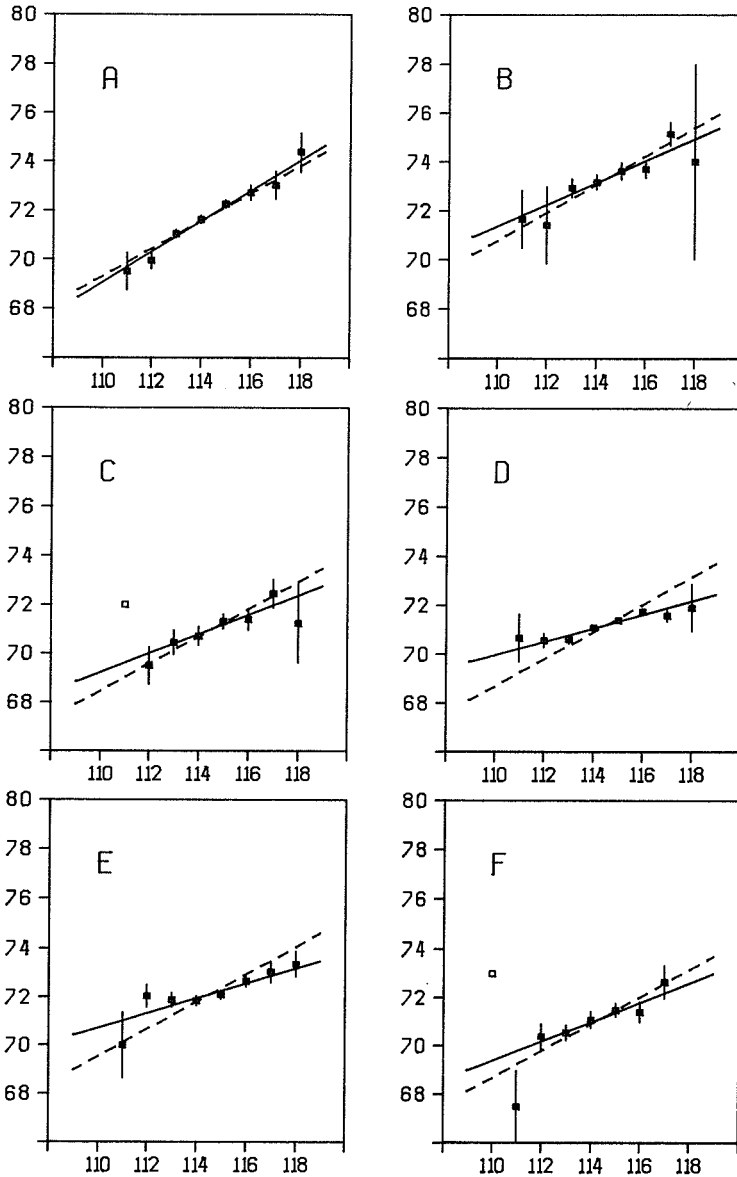
Fig. 1. A-F. Data for Højer subsamples: mean length and S.D. of the mean for each TNV (filled squares denote mean length, vertical bars extend to ± 1 standard deviation of the mean; open squares denote single specimens). Full line: linear regression of length on TNV. Broken line: theoretical dependence of length on TNV under the 'proportionality' hypothesis. Abscissa for each graph: TNV. Ordinate for each graph: length (mm).

Graphs correspond to samples and developmental stages as follows:

A: Sample I, Stage A   C: Sample II, Stage A   E: Sample III, Stage C
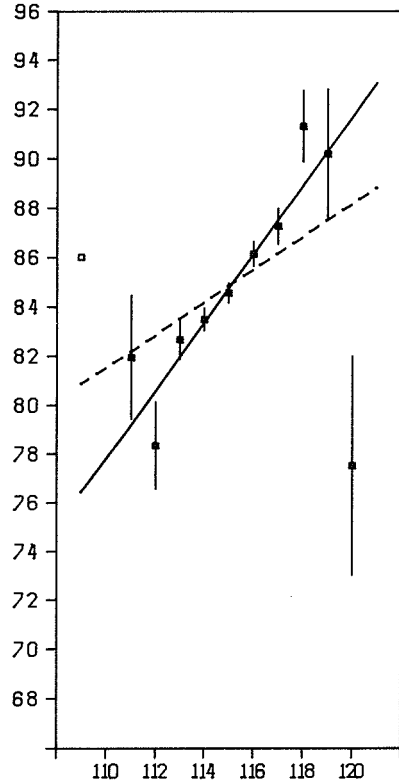B: Sample I, Stage D   D: Sample II, Stage D   F: Sample III, Stage E

Fig. 2. Data for Esrom sample.
Interpretation of the graph as for Fig. 1.

Boëtius (1976, fig. 8) shows that in the Esrom sample there is a similar dependence of mean TNV on length, slightly less variable. In Fig. 2 we present the dependence of mean length on TNV, and a clear positive relationship, markedly stronger than for Højer, is seen.

Table 3 gives summary statistics for the Højer sub-samples and the Esrom sample.

Table 3. Højer material, divided into sub-samples by developmental stage, and Esrom material: slope of regression of length on TNV (with S.D.), mean length, mean TNV, ratio of these and difference (Devn) between observed and theoretical slopes (using slope = $0.9 \times \bar{L}/\bar{V}$) as multiple of S.D. of this difference.

Sample	Slope	S.D.	$\bar{L}$	$\bar{V}$	$\bar{L}/\bar{V}$	Devn
IA:	0.622	0.073	71.70	114.21	0.628	+0.78
ID:	0.445	0.137	73.42	114.57	0.641	−0.97
IIA:	0.395	0.136	71.07	114.65	0.620	−1.20
IID:	0.280	0.050	71.24	114.55	0.622	−5.61
IIIC:	0.308	0.084	72.16	114.63	0.630	−3.08
IIIE:	0.403	0.133	71.21	114.49	0.622	−1.18
Esrom:	1.387	0.172	84.94	115.13	0.738	+4.20

Hypotheses about the TNV- Length relationship

(A) One may posit that the dependence of mean length on TNV is a matter of simple proportionality: the more segments make up an individual, the greater *pro rata* its expected length. This could be in keeping with the data for Højer, where the slopes for different sub-samples are similar. However, the proportionality clearly breaks down for Esrom, where the slope is very different. The following considerations are also relevant:

(i) On this hypothesis, it would be the segment rather than the individual whole organism that was the 'unit of growth'.

(ii) If a 'random' vertebra were added or subtracted, a theoretical dependence of average length on TNV can be calculated. This is done below, where it is shown that the theoretical slope is generally greater than the observed slope for Højer (and less, for Esrom). Also, there are statistically significant differences between the slopes for Højer sub-samples.

(iii) This mechanism would be mostly compatible with the Højer data, if the variation in vertebral complement was associated with the shorter vertebrae. On the other hand, the mechanism would then be an even worse match to the Esrom data.

(B) Along the lines suggested for, for instance, North Sea herring, 'individuals with a higher genetic potential for growth tend to have higher numbers of vertebrae' (MAFF 1967, p.34). This could be compatible with homogeneity if the elver population has a continuous cline in this genetic potential. Otherwise (if the 'cline' is discontinuous) Schmidt's homogeneity hypothesis is *ipso facto* contradicted.

(C) Numbers of vertebrae increase as growth proceeds, even through the stages of elver and small yellow eel.

(i) This would seem unlikely for so definite a structure as the vertebrae. The progressive differentiation of myomeres in small (<30 mm) larvae is another matter, but occurs at a much earlier stage.

(ii) The relationship is nearly independent of developmental stage and of time of year. Also, there is no uniform relationship between mean TNV and developmental stage or time of year (Boëtius 1976, table 4 – especially when specimens with TNV$\leqq$110 are not counted), nor between mean length and developmental stage or time of year (loc. cit. Table 2); and there is no relationship between mean length and mean TNV for sub-samples (Table 4). See below for further discussion.

Table 4. Mean TNV and mean Length for the Højer material according to developmental stage.

	IA	ID	IIA	IID	IIIC	IIIE
Mean TNV = 114+	0.209	0.566	0.646	0.533	0.630	0.473
Mean Length	72.16	73.92	71.56	71.73	72.64	71.67

(D) Application of Schmidt's principle for counting TNV (in particular the determination of the 'last hour-glass shaped vertebra') may sometimes give different results according to length, as a result of morphological changes associated with growth.

(E) The Højer specimens are a combination ('mixture') of distinct groups, where in each group there is no relationship between TNV and length (i.e. mean length is independent of TNV), but the groups differ between themselves in mean TNV and also in mean length. It is not suggested that such differences between groups would necessarily reflect genetic inhomogeneity, nor that they would correspond to different spawning-grounds: if this hypothesis is true, the differences could arise through differing environmental conditions in earlier development associated with different routes to the Danish coast. The hypothesis will be analysed in detail below.

Theoretical analysis of 'random missing vertebrae'

With reference to Hypothesis (A), suppose that TNV for an individual differs from a standard TNV by insertion or deletion of a random vertebra. The length of an elver's vertebral column is about 90% of the elver's total length, so the average length of a vertebra of an elver of length L and TNV$= V$ is $0.9 \times L/V$. Length of vertebral column as fraction of total length may be somewhat variable, but we can be fairly sure that it is always at least 85%. Ege (1939) finds that the length of the 'head' (tip of snout to gill opening) varies between 11% and 14% of total body length; this is about twice the distance from the snout to the beginning of the vertebral column. To this 5-7% must be added some 3-5% for the distance from the last vertebra to the tip of the tail.

If F is the ratio (vertebral column)/(body length), then the slope of mean length $\bar{L}$ versus TNV V should be

$$\text{Slope} = F \cdot \bar{L}/V$$

under Hypothesis (A). From Table 3 it is seen that four out of the six sub-samples from Højer differ by about one S.D. from the prediction with $F = 0.9$, the other two (IID and IIIC) being at least 3 S.D.s away. Apart from IA, all are lower than predicted.

The weighted sample variance of all six slopes (weight $= 1/(\text{S.D.}^2)$ is 16.06, to be compared with χ^2 on 5 degrees of freedom ($P = 0.006$). Thus the six slopes are significantly different. Various subsets of them can be found which do not differ significantly, e.g.

(i) ID, IIA, IIIC, IIIE ($\chi^2 = 0.92$ on 3 d.f., $P = 0.8$)
(ii) IA, ID, IIA, IIIE ($\chi^2 = 3.89$ on 3 d.f., $P = 0.27$)

Similar data for the Esrom sample are also given in Table 3. The observed slope is significantly (more than 4 S.D.s) greater than the predicted.

These results show that the 'proportionality' hypothesis (A) apparently does not hold for the Esrom sample. While it may hold for a majority of the Højer sub-
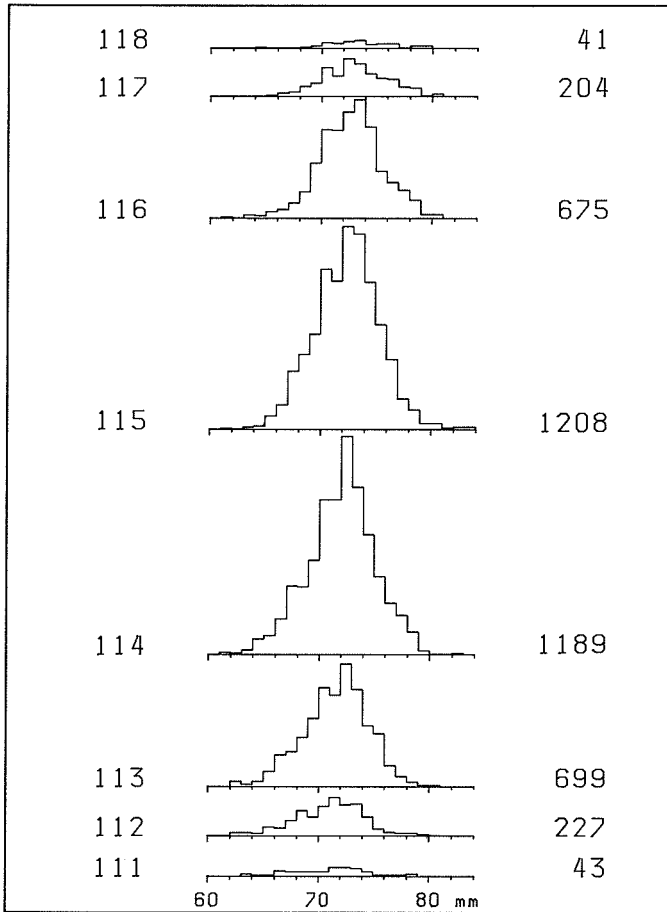
Fig. 3. Histograms of length distribution for each TNV, for the Højer
material (all three samples pooled). On the left: TNV. On the right:
observed numbers of specimens for each TNV. Heights of histogram
bars are proportional to absolute numbers. Histogram for TNV = 119
(3 specimens at 71, 76 and 78 mm) not shown.

samples, it certainly does not do so uniformly. The hypothesis cannot be rejected
outright, but a search for alternatives is indicated. Of those listed above, the only
one that can be tested in the data is (E), the 'mixture' hypothesis.

The mixture hypothesis

The biological interpretation motivating the 'mixture' hypothesis for the Højer
data is that, for a group of elvers of common origin and at the same stage of
development, the distribution of length (L) would be independent of numbers of
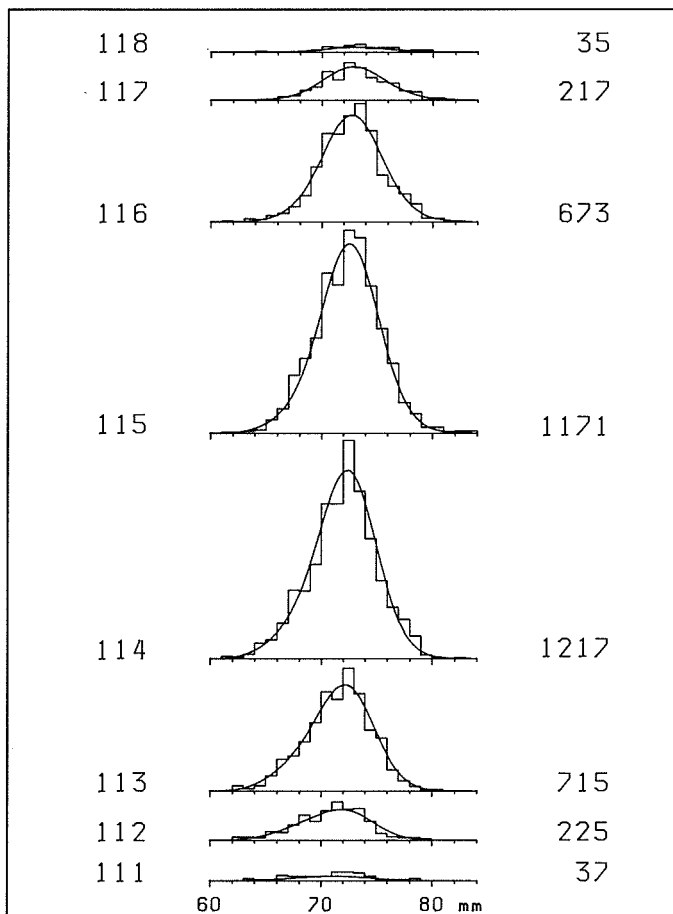vertebrae (TNV). This is consistent with material in which average length mani-

Fig. 4. The fitted 'mixture' model superimposed on the histograms of Fig. 3. On the left: TNV. On the right: expected number (E_V) calculated from the fit as explained in the text.

festly varies according to TNV, provided the sample consists of two or more such groups, in each of which L is independent of TNV, and the different groups are characterised by (a) different TNV distributions and (b) different length distributions. The relationship between mean length and TNV can be very closely linear.

Fig. 3 presents the Højer data on L and TNV for all samples pooled, as histograms of L for each TNV (TNV = 119, with 3 specimens only, is not shown). Inspection of the figure indicates that

(i) there is a central peak which is in a constant position for TNV = 113-115 (mean approx. = 72.5 mm);
(ii) the length distributions for TNV = 112-114 are negatively skew;
(iii) the length distributions for TNV = 116-118 are positively skew.

Note also that there is indication, especially near the peaks, of a preference for recording certain lengths rather than others. Thus lengths of 70 mm and 72 mm and possibly 75 mm appear disproportionately frequent, lengths of 69 mm and 71 mm and possibly 74 mm disproportionately infrequent, which can be seen also from Primary Table II of Boëtius (1976). This kind of artefactual recording error is in fact very common when large numbers of items are being measured routinely (see for instance Cunliffe (1976)). It normally has a minor influence on calculated quantities such as means and variances, and is mentioned specially only because of its somewhat disruptive effect on the rather delicate model-fitting, which depends on fine detail of the shapes of distributions.

The model adopted for the data is based on the above considerations: the large fixed central peak is one component, with mean TNV and mean L both having near-central values. The negative skewness for lower TNV is attributed to a minor component where mean TNV and mean L both have relatively low values, and the positive skewness for higher TNV is attributed to a minor component where mean TNV and mean L both have relatively high values. There is substantial overlap between the central component and each of the other two.

Thus the model for the data can be written as an expression for the probability $P(V, L)$ that a specimen will have TNV $= V$ and length $= L$, in the form

$$P(V, L) = \pi_1 P_1(V) \cdot Q_1(L) + \pi_2 P_2(V) \cdot Q_2(L) + \pi_3 P_3(V) \cdot Q_3(L)$$

where

π_1, π_2 and π_3 are the proportions of the three components (adding to 1).

$P_i(V)$ $(i = 1, 2, 3)$ is the probability, calculated from a normal distribution with mean MV_i and standard deviation SV_i, that a sample from that distribution will lie between V and $V + 1$ (i.e. the integral of this distribution from V to $V + 1$). This amounts to modelling TNV as if it were the discrete expression of an underlying continuous variable, the latter being supposed to have the normal distribution, $Q_i(L)$ $(i = 1, 2, 3)$ is the same, calculated from a normal distribution with mean ML_i and standard deviation SL_i, for the event of observing a length between L and $L + 1$,

and

$$MV_1 < MV_2 < MV_3 \text{ and } ML_1 < ML_2 < ML_3.$$

There are therefore 14 parameters. Their values were estimated by an iterative procedure based on the Maximum Likelihood approach. The independence of TNV and L in each component group is assured, since their joint distribution in each component is the product of their separate distributions.

Fitting the model

Reasonable initial estimates were obtained by a preliminary analysis, making some simplifying assumptions. It was assumed that all the SV_i were equal, and all the SL_i. The SV_i were set at the value 1.1, typical of the more compact distributions of vertebra counts obtained in the past (see e.g. Boëtius 1980). It was supposed that

the histogram for TNV = 113 (699 specimens) was almost entirely groups 1 and 2, with an unknown proportion p of group 1. In such a distribution the counts for the lowest values of L will arise almost purely from group 1; if the number in group 1 were known, a normal probability plot could be made from which the mean and S.D. of L could be estimated. This was approximated by varying p until the probability plot was initially a straight line. The same was done with the largest values of L in the histograms for TNV = 116 and 117. Thus groups 1 and 3 received a preliminary estimate. The mean of L for group 2 was taken as 72.5, and its S.D. as equal to that estimated for groups 1 and 3. Approximate values for MV_1, MV_2 and MV_3 were obtained by consideration of 'components of variance', using the assumed $SV_i = 1.1$, and finally the proportions were approximately estimated by matching theoretical dependence of L on V with observed.

Starting with these initial estimates, an iterative procedure was entered where the parameters were varied in such a way as to increase the likelihood in each iteration. This was done interactively, and care was taken to arrest iterative parameter variation when the resulting fitted distributions began to be unduly influenced by potentially artefactual features of the data, as described above.

Consequently, the parameter values finally chosen are not the maximum likelihood estimates, since iteration was stopped before the maximum was attained. For this reason, estimates of standard error of the parameter-estimates cannot be obtained by variation of the likelihood function. The overall fit is, however, satisfactory.

Table 5. Parameters of the three-component mixture model fitted to the Højer data. For each component: estimated proportion of the whole material, mean and standard deviation of the fitted distribution of TNV, mean and standard deviation of the fitted distribution of length.

Component 1: Proportion = 17.3 % (742 specimens)		
TNV:	Mean = 114.20	S.D. = 1.16
Length:	Mean = 68.96	S.D. = 2.77
Component 2: Proportion = 62.0 % (2659 specimens)		
TNV:	Mean = 114.80	S.D. = 1.14
Length:	Mean = 72.64	S.D. = 2.42
Component 3: Proportion = 20.7 % (888 specimens)		
TNV:	Mean = 116.13	S.D. = 1.03
Length:	Mean = 73.27	S.D. = 2.94

The parameter values finally adopted are presented in Table 5, and Fig. 4 shows the resulting mixed expected distributions superimposed on the original histograms. The expected number E_V for each value of TNV was calculated as

$$E_V = 4289(\pi_1 P_1(V) + \pi_2 P_2(V) + \pi_3 P_3(V)),$$

i.e. no attempt was made to force the expected marginal totals for TNV or for L to agree with observed marginal totals.

Another possible representation of the data could be a mixture of two groups only. One group would be the invariable dominant central group as above; the other would be a group in which TNV and L co-vary as described in Hypotheses (A), (B) or (C) above (i.e. in this group there would be dependence between TNV and L). The second group would be largely masked by the first for TNV = 114 or 115, but would appear on the left for TNV = 111-113, and on the right for TNV = 116-118. Such a model has not been explicitly fitted.

Discussion to Part 2

The work just described has shown that it is possible to represent the Højer data as a mixture of three components, with different distributions of length and different distributions of TNV. If this corresponds to three really distinct groups of elvers, then the Højer elvers are not homogeneous. The presence of three different length-groups may reflect no more than the presence of different 'waves' of ascent; the associated presence of three TNV-groups is more fundamental, and indicates heterogeneity, unless a 'joint growth' hypothesis such as (A), (B) or (C) above is

Table 6. Højer material divided into sub-samples by developmental stage. For each TNV, observed numbers (V), numbers expected (E) on the hypothesis that TNV distribution does not depend on sub-sample, and deviation from expectation (D: see text). Højer sample IIE (15 specimens only) omitted.

Sample:		IA	ID	IIA	IID	IIIC	IIIE
111	V:	16	3	2	13	7	2
	E:	10.3	3.0	2.6	18.8	5.4	3.0
	D:	+1.8	0	−0.4	−1.3	+0.7	−0.6
112	V:	76	7	14	98	20	12
	E:	54.3	15.7	13.6	99.1	28.7	15.6
	D:	+2.9	−2.2	+0.1	−0.1	−1.6	−0.9
113	V:	214	49	34	274	72	53
	E:	166.6	48.0	41.8	303.8	88.1	47.7
	D:	+3.7	+0.1	−1.2	−1.7	−1.7	+0.8
114	V:	303	85	65	506	151	77
	E:	284.0	81.9	71.4	518.1	150.2	81.4
	D:	+1.1	+0.3	−0.8	−0.5	+0.1	−0.5
115	V:	255	84	77	543	156	89
	E:	288.1	83.1	72.4	525.5	152.4	82.5
	D:	−2.0	+0.1	+0.5	+0.8	+0.3	+0.7
116	V:	113	50	42	321	100	44
	E:	160.3	46.2	40.3	292.4	84.8	45.9
	D:	−3.7	+0.6	+0.3	+1.7	+1.7	−0.3
117	V:	37	14	19	96	23	15
	E:	48.8	14.1	12.3	89.0	25.8	14.0
	D:	−1.7	−0.3	+1.9	+0.7	−0.6	+0.3
118	V:	9	2	4	14	11	1
	E:	9.8	2.8	2.5	17.9	5.2	2.8
	D:	−0.3	−0.5	+1.0	−0.9	+2.6	−1.1
119	V:	0	1	0	1	1	0

accepted. Such heterogeneity need not correspond to three genetically distinct spawning populations, nor even to distinct spawning grounds.

All three groups, if real, must be present in all six sub-samples, though in slightly different proportions. If the apparently tripartite nature of the pooled material is accepted as the explanation of the length-TNV relationship, it is not merely a consequence of the fact that there are simply three different samples taken at different times, since the length-TNV relationship is equally manifest in all sub-samples.

Table 6 shows the relationship between observed vertebra counts (V), expected vertebra counts (E) on the assumption that TNV distribution is independent of sub-sample, and sub-sample. Also shown is a measure of deviation from expectation:

$$D = (V - E)/\sqrt{E}.$$

It can be seen that, in terms of the three groups 1, 2 and 3 in the hypothesised tripartite representation, there is excess of group 1 in sub-sample IA, and excess of group 3 in sub-samples IIA, IID and possibly IIIC. On this interpretation, group 1 (with a low mean TNV) occurs more at the beginning of the elver run (and for developmental stage A rather than D), and group 3 (with a very high mean TNV) more in the middle and towards the end of the elver run (and nearly independently of developmental stage). Group 2 is generally present throughout.

Final conclusions

The homogeneity claimed by Schmidt (1912 and 1913) for the entire European population of *Anguilla anguilla*, based on counts of total numbers of vertebrae (TNV) in several samples from widely spread locations, has been tested in three ways.

First, analysis of Schmidt's published results shows that, in a test for excessive uniformity, they are unusually homogeneous.

Second, analysis of data first published by Boëtius (1980), being partly results obtained by Schmidt but not previously published, and partly results obtained by Boëtius from preserved material left by Schmidt, shows that these results are not homogeneous. In the Atlantic a north-south cline of increasing TNV is observed. In the Mediterranean, the results are significantly heterogeneous, but there is no apparent geographical cline. These results admit, separately, various explanations. Taken together, while ambiguous, they tend to show that Schmidt's published results were excessively uniform.

Third, large samples of elvers and small eels from Højer and Esrom in Denmark were analysed. These samples exhibit a clear relationship between length and TNV. A 'proportionality' hypothesis, whereby length varies *pro rata* with TNV, roughly matches the Højer data, but is not uniformly consistent with sub-samples (which also differ significantly among themselves), and is quite inconsistent with the Esrom data. It is shown that the Højer samples can be described in terms of a mixture of three distinct groups, with different TNV distributions and different length distributions. In each group, length is independent of TNV. These three groups, if real, are present in all samples and at all stages of development, though

in somewhat different proportions. If this model is accepted, the elver run at Højer is not homogeneous.

Possible explanations of such inhomogeneity include difference in environmental conditions during very early development at different parts of the one spawning area or different environmental conditions encountered during movement to the coasts along different routes, as well as a widely dispersed spawning area or even several disjoint spawning areas. In any case, Schmidt's claim appears weakened.

Acknowledgments

References

Boëtius, J., 1976: Elvers, Anguilla anguilla and Anguilla rostrata from two Danish localities. Size, body weight, developmental stage and number of vertebrae related to time of ascent. – Meddr Danm. Fisk.- og Havunders. N.S. 7: 199-220.

Boëtius, J., 1980: Atlantic Anguilla. A presentation of old and new data of total numbers of vertebrae with special reference to the occurrence of Anguilla rostrata in Europe. – Dana 1: 93-112.

Cunliffe, S.V., 1976: Interaction. – J.R. Statist. Soc. A., 139: 1-19.

Ege, V., 1939: A revision of the genus Anguilla Shaw. A systematic, phylogenetical and geographical study. – Dana-Report No. 16.

MAFF, 1967: Annual Report of the Director of Fisheries Research, 1967 (Ministry of Agriculture, Fisheries and Food), Lowestoft.

Schmidt, Johs., 1912: Danish Researches in the Atlantic and Mediterranean on the Life-History of the Freshwater Eel (Anguilla vulgaris Turt.). – Intern. Rev. der ges. Hydrobiol. und Hydrographie, Leipzig, Nov. 1912: 317-342.

Schmidt, Johs., 1913: First Report on Eel Investigations 1913. – Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer. 18: 1-29.

Schmidt, Johs., 1917-1930 (ed.): Racial Investigations I-IX, 1917-1930. – Comptes rendus des travaux du Laboratoire Carlsberg, Vols 13-18.