AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 79:117–120(1989)

Test of Phenice's Technique for Determining Sex From the Os Pubis

NANCY C. LOVELL
Department of Anthropology, Cornell University, Ithaca, New York 14853

KEY WORDS Forensic Anthropology, Pubic bone morphology, Osteological analysis

ABSTRACT Pubic bone morphology was examined to test the accuracy of Phenice's visual method for determining sex from the os pubis. Twelve participants scored 50 pubic bones from individuals of known sex aged 52–92 years. The sample is of modern males and females, all presumed whites. An accuracy of ~83% in determining sex was recorded, compared to 95% reported by Phenice. This accuracy difference may reflect different age distributions of the two samples. Through replication of test results on two series of 25 specimens, the technique was found to be reliable. Previous experience in human osteological analysis was shown to have no effect on accuracy in this test, confirming Phenice's assertion that the technique does not require extensive experience to yield accurate results. Results suggest that there is a moderate negative correlation between accuracy in determining the sex of an individual and that individual's age.

A visual method for determining sex from the os pubis was developed by Phenice (1969), in which three morphological features–ventral arc, subpubic concavity, and medial aspect of the ischiopubic ramus–were reported to be very useful indicators of sex, providing greater than 95% accuracy. Phenice's technique is widely used in human osteological analysis because of its relative ease of application and apparent degree of accuracy (for example, see Bass, 1971; Ortner and Putschar, 1981; Skinner and Lazenby, 1983; Stewart, 1979; Ubelaker, 1978). However, the validity of the technique has not yet been firmly established, since Kelley's (1978) critique of the method is based on prehistoric skeletal remains of unknown sex and age, and Sutherland and Suchey's (1987) analysis of the ventral arc criterion does not include individuals over 30 years of age.

The purpose of the present study is to apply the method to a sample with age and sex known for each specimen in order to establish firmly the reliability of the technique. This test evaluates accuracy as a function of long-term and short-term observer training and the effects of sex and age distribution of the skeletal sample.

MATERIALS AND METHODS

Fifty bone specimens, representing 36 presumed white individuals (13 males and 23 females), were randomly selected from a larger series of pubic bones obtained from medical school cadavers. The bones had been sawn through the superior ramus of the os pubis and the ischiopubic ramus, thereby preserving the pubic symphyseal region. This method is commonly used to remove the pubic elements for forensic analysis. These partial innominates reduced the likelihood of participants using other sex-related features, such as sciatic notch morphology, to aid identification.

Twelve participants, undergraduate and graduate students and a professional physical anthropologist, completed the test. They were asked to rate their experience in human osteological analysis as none (they had not examined human or nonhuman skeletal remains before), moderate (they were enrolled in an undergraduate human osteology

Received August 28, 1987; accepted June 15, 1988.

Nancy C. Lovell's present address is Department of Anthropology, 13-15 Tory Building, University of Alberta, Edmonton, Alta. T6G 2H4, Canada.

TABLE 1. Results of sex determination test

Number of correct sex determinations (each series size = 25)		Observer's experience in osteological analysis
Series A	Series B	
18	19	None
19	18	Moderate
19	23	Considerable
20	18	Considerable
20	20	None
21	19	None
21	23	Moderate
22	20	None
22	20	Considerable
22	22	Considerable
23	21	Moderate
24	22	Considerable
20.9 (mean)[1]	20.4 (mean)[1]	

[1]Not a statistically significant difference.

course), or considerable (graduate students conducting research on human skeletal remains and the professional physical anthropologist). All received the same instruction in the application of the technique: Phenice's (1969) article was verbally summarized, and the presence and absence of the three features were indicated on Phenice's diagram (1969: 299), on male and female articulated pelves from study collections, and on demonstration samples similar to the specimens used in the test. Participants were permitted to refer back to the diagram and demonstration samples during the test.

The bone specimens were randomly separated into two series of 25 samples each and arranged on a table. Participants were asked to determine the sex of each pubic bone in series A, followed by series B, on the basis of Phenice's three criteria. These criteria are the presence or absence of a ventral arc, the contour of the subpubic area, and the shape of the medial aspect of the ischiopubic ramus. Participants were able to relax between series identifications but did not receive further training, nor were they told their results from the first series. Thus "short-term training" refers only to gains in familiarity with the specimens as a result of examining the first series.

Since the autopsy method of removing the os pubis may result in the specimens being sawn closely to the symphyseal rim, evaluation of the contour of the subpubic concavity may be difficult. To assess this factor, participants scored the presence or absence of each criterion, then assigned the sex for the individual, using Phenice's recommendations for weighting the criteria and were also asked to comment on the ease of application of each of the criteria and whether the bone was cut too closely for evaluation.

RESULTS

In Table 1, accuracy is expressed as the number of specimens, out of 25 (per series), that were correctly sexed. The third column indicates participants' experience. Average accuracy is 83% ± 7%. The highest accuracy obtained was 92%, by the professional physical anthropologist.

To evaluate the association between participants' previous experience in human osteological analysis and their accuracy in this test, a one-way analysis of variance was performed on combined mean scores for participants in each of the three levels of expertise as described in Table 2. In spite of the fact that the professional physical anthropologist scored highest in accuracy, the difference among the accuracies obtained by these groups is not statistically significant at alpha = .05.

The effect of short-term training was tested by the replication of results, using the split-halves method (i.e., series A and B). Halves are considered approximations to alternate forms of the same test and have the advantage of being conducted on one occasion. To determine if practice resulted in improvement of accuracy, a paired t test was used to compare scores on the halves. There was no significant change in scores at alpha = .05. This is also true if the scores are partitioned and examined for each level of experience. The reliability of the method was determined by computing a correlation coefficient, using the Spearman-Brown prophecy formula (Carmines and Zeller, 1979). Here, R = .55, indicating acceptable reliability.

The majority of participants commented that the ventral arc caused the most uncertainty because of the often highly irregular surfaces on bones of aged individuals, which

TABLE 2. Participants' accuracy vs. their experience in osteological analysis

Experience	Mean accuracy	Percentage accuracy
None	20.8[1]	83[2]
Moderate	20.3[1]	81[2]
Considerable	21.0[1]	84[2]

[1]Not statistically significant differences.
[2]Not statistically significant differences.

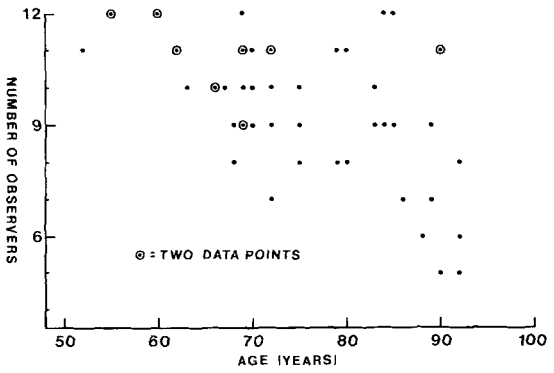

Fig. 1. Number of observers making correct sex determinations vs. age at death for specimen.

sometimes masked expression of the feature. This finding contrasts with other studies, in which the ventral arc is the least ambiguous feature (Kelley, 1978; Phenice, 1969; Sutherland and Suchey, 1987). In spite of this perceived difference in usefulness of these features, there was no differential success in the application of the three criteria in this test (ANOVA at alpha = .05). For 45 of the 600 sex determinations, it was noted that the bone was cut short enough to make evaluation of the subpubic concavity difficult; however, in only three of these cases was the sex assigned incorrectly.

In the evaluation of sample characteristics, it was determined that one sex was not easier to identify than the other. Eighteen percent of the male specimens were incorrectly sexed, and 17% of the female specimens were incorrectly sexed. The difference is not significant by χ^2 analysis at alpha = .05. This supports Kelley's (1978) conclusion that, although female specimens may be more ambiguous, relative accuracy is not affected.

The second sample characteristic evaluated was age distribution. The relationship between the age at death for each individual specimen, and the accuracy obtained in determining its sex in this test is shown in Figure 1. The vertical axis is the number of observers, out of 12, who correctly sexed the specimen, and the horizontal axis is the age at death of the individual. A correlation statistic ($r = -.48$) suggests that accuracy may decrease with the age at death of the individual being examined. In addition, there is increased variability in accuracy obtained for older ages.

DISCUSSION

Phenice's (1969) accuracy is clearly out of the range of the results obtained in this study. A comparison of the average accuracies, taking into account the variance of this accuracy estimate and the unequal sample sizes, indicates that the difference is statistically significant at the .05 level. This decreased accuracy, and its large associated variation, could cast some doubt on the validity of Phenice's technique. Although it has been shown that accuracy, in this study, is not a function of long-term or short-term observer training, nor of the sex distribution of the sample, it appears that age may be a factor. The difference in accuracy between these and Phenice's results may represent different age distributions of the two samples. Two hundred and seventy-five individuals from the Terry Collection were used in the original study (Phenice, 1969: 298); however, the age data were not published, so the age range of that sample is unknown. Although average age varies for Terry Collection subcategories, such as males and females, or blacks and whites (see, for example, Angel, 1982), it is reasonable to assume an average age of 50–55 years for individuals in the collection (Angel, personal communication) and to extrapolate that average age to Phenice's sample. This is much younger than the mean age for the sample used in this test, which suggests that the age difference is responsible for the accuracy difference. This suggestion is not unreasonable if we consider that degenerative bony changes associated with advanced age may mask the expression of one or more of Phenice's diagnostic features.

CONCLUSIONS

An accuracy in determining sex of ~83% was obtained in this study, compared to 95% reported by Phenice. The technique was found to be reliable by replication of results, and its accuracy not affected by the observer's previous experience in osteological analysis. However, accuracy appears to decrease when the method is used on older individuals. This trend may explain the differences in average accuracies obtained in this and the original study, which was apparently based on younger individuals.

ACKNOWLEDGMENTS

I thank Mark Skinner, Richard Lazenby, and the students at Simon Fraser University

who participated in this study. Specimens were collected and contributed to S.F.U. by Owen Beattie and Carlos Germann. J. Lawrence Angel provided access to Terry Collection records. Kenneth A.R. Kennedy, Jere Haas, and Mark Skinner gave valuable criticisms of the manuscript. Funding was provided by the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Graduate School, Cornell University.

LITERATURE CITED

Angel JL (1982) A new measure of growth efficiency: skull base height. Am. J. Phys. Anthropol. *58:*297–305.

Bass WM (1971) Human Osteology: A Laboratory and Field Manual, 2nd ed. Columbia, MO: Missouri Archeology Society.

Carmines EG, and Zeller FA (1979) Reliability and Validity Assessment. Beverly Hills, CA: Sage.

Kelley MA (1978) Phenice's visual sexing technique for the os pubis: a critique. Am. J. Phys. Anthropol. *48:*121–122.

Ortner DJ, and Putschar WGJ (1981) Identification of Pathological Conditions in Human Skeletal Remains. Washington, DC: Smithsonian Institution Press (Contributions to Anthropology No. 28).

Phenice TW (1969) A newly developed visual method of sexing of the os pubis. Am. J. Phys. Anthropol. *30:*297–302.

Skinner M, and Lazenby RA (1983) Found! Human Remains. Burnaby, BC: Archaeology Press Simon Fraser University.

Stewart TD (1979) Essentials of Forensic Anthropology. Springfield, IL: Charles C. Thomas.

Sutherland LD, and Suchey JM (1987) Use of the ventral arc in sex determination of the os pubis. Poster presentation, 39th Annual Meeting of the American Academy of Forensic Sciences, Physical Anthropology Section, San Diego, CA.

Ubelaker DH (1978) Human Skeletal Remains: Excavation, Analysis, Interpretation. Washington, DC: Taraxacum Manuals on Archaeology No. 2.