This is to certify that the

thesis entitled

COMPARISON OF AN ESTIMATE OF SKELETAL AGE
WITH CHRONOLOGICAL AGE
WHEN CLASSIFYING ADOLESCENT MALES
FOR MOTOR PROFICIENCY NORMS

presented by

Michael Grant Marshall

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Health, Physical Education, and Recreation

Major professor

Date April 19, 1978

O-7639

COMPARISON OF AN ESTIMATE OF SKELETAL AGE

WITH CHRONOLOGICAL AGE

WHEN CLASSIFYING ADOLESCENT MALES

FOR MOTOR PROFICIENCY NORMS

By

Michael Grant Marshall

A THESIS

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

DOCTOR OF PHILOSOPHY

Department of Health, Physical Education
and Recreation

1978

ABSTRACT

COMPARISON OF AN ESTIMATE OF SKELETAL AGE WITH CHRONOLOGICAL AGE WHEN CLASSIFYING ADOLESCENT MALES FOR MOTOR PROFICIENCY NORMS

By

Michael Grant Marshall

This investigation sought 1) to determine the proportion of the variance in the skeletal development of adolescent males that is explained by chronological age, height, weight, and a Sexual Maturation Indicator Value (SMIV) estimate, and 2) to determine how an estimate of skeletal age (ESA) that incorporated chronological age, height, weight, and SMIV compared with chronological age when classifying adolescent males for motor proficiency norms.

To answer part one required a sample of adolescent males who represented a rectangular distribution of sexual maturation. Hand-wrist X-rays, height and weight were obtained on 30 randomly selected boys from East Lansing, Michigan who ranged in age from 134.5 to 180.0 months. Four Physical Education teachers and the author assessed five primary and secondary sexual characteristics on two separate occasions. The sexual characteristics were; 1) facial hair, 2) axillary hair, 3) pubic hair, 4) penis development and 5) scrotum and testes development. Via the Greulich-Pyle Atlas, an experienced rater of hand-wrist X-rays assessed the skeletal ages of the boys on two separate occasions.

The analysis determined that for this sample: (1) chronological age explained 45 percent of the variance in the skeletal age of the adolescent males; (2) height explained an additional 20 percent; (3) weight contributed an additional 6 percent; and (4) the Sexual

Maturation Indicator Value (SMIV) explained 12 percent more. Therefore, the combination of chronological age, height, weight and an estimate of sexual maturity explained 83 percent of the variance in the skeletal age of the sample.

The regression analysis indicated that the most efficient combination of secondary sex characteristics for estimating SMIV was:

$$SMIV = 126.47 + 174(X(F)) + 1.20(X(A)) + 2.75(X(PH)) + 6.68(X(P))$$

where; F is the boys' facial hair rating, A is the boys' axillary hair rating, PH is the boys' pubic hair rating, and P is the boys' rating of penis development.

To answer part two required a large cross-section sample of adolescent males. Therefore, at the author's request, the boys' physical education teachers in the East Lansing and Mason, Michigan Junior High Schools collected the required data on the entire male populations of their schools for three years. The cross-sectional sample consisted of 1578 boys on whom their teachers determined their chronological age, measured their height and weight, tested their ability to perform sit-ups, pull-ups, standing long jumps, vertical jumps, and shuttle runs, and assessed five primary and secondary sexual characteristics.

The Multivariate Analysis of Variance computer program calculated the raw regression coefficient and mean values for the chronological age, height, weight, and SMIV estimate variables to create the estimate of skeletal age (ESA) regression equation. By using identical classification intervals for the ESA estimate and the chronological age (CA) value, the analysis examined the within-cell variances for the five motor proficiency tests and the mean value progressions across the annual increments.

Although the within cell variances for the classifications by CA and ESA were not significantly different, the author concluded that compared to CA, ESA classified adolescent males for motor proficiency tests better because: 1) ESA is less likely to reward the competitors of the basis of advanced biological maturation; and 2) whereas CA had a preponderance of 13 and 14 year old subjects, SA dispersed the subjects throughout the adolescent age range, indicating that ESA moved the accelerated and delayed-maturers to groupings of similar biological maturation. The regression analysis indicated that the most efficient estimate of skeletal age was:

$$ESA = -16.15 + .277(X(CA)) + .212(X(H)) + .220(X(W)) = .594(X(SMIV))$$

where; CA is the boys' chronological age in months, H is the boy's height in inches, W is the boys' weight in pounds, and SMIV is the estimate of the boys' sexual maturation.

DEDICATION:

To Moose, Runt, and Pizwilly, three of the greatest kids that a Dad could have.

And, to all of the boys whose athletic desires the present system of athletic competition vanquished.

ACKNOWLEDGEMENTS

At the East Lansing and Mason, Michigan Junior High Schools, the author thanks the Boys Physical Education teachers; Craig Marsh, Wally Juall, Ron Hornung, Dave Wills, Rich Wilson, and Marvin Pulver. Special thanks goes to Jim Oestriech for his ideas, help, and encouragement, to Sal DiFranco for his interest, and to the boys.

At Michigan State University, the author thanks his doctoral guidance committee; chairman Vern Seefeldt, Glenn Hatton, Bill Heusner, Rex Carrow, and Bill Schmidt. For his many hours of extra help, special thanks goes to Bill Schmidt.

TABLE OF CONTENTS

APPENDICES Page

APPENDIX

LIST OF TABLES

Table		Page

Table		Page

2.

2.

2.

2.

2.

2.

2.

2.

2.

2.

2.

2.

2.

Table		Page

Table | Page

Table		Page

LIST OF FIGURES

Figure		Page

CHAPTER I

INTRODUCTION

In the United States, competition is a way of life. In the classroom, teachers use competition for grades to motivate students to improve their academic skills. On the athletic fields, coaches use competition for victory to motivate athletes to improve their motor skills. Supporting the value of competition in education is the principle that regardless of the competition's outcome, if the students put forth their greatest effort, then everybody wins. However, for competition to remain an ongoing motivation, it must reward the quality of effort and not factors over which the competitors have no control. In athletic competitions for adolescent males, the participants have no control over their biological maturation. Therefore, to reduce the influence that biological maturation has on the outcome of the competition and to increase the likelihood that the rewards match the quality of effort, the sponsors of athletic competitions for adolescent males should: (1) equate the competitors according to biological age, or (2) adjust the rules to eliminate the advantage that early-maturing boys have over their later-maturing peers.

When teaching physical education classes in East Lansing Junior High School, the author encountered an athletic competition for adolescent males for which he suspected that nobody had equated the competitor's biological maturation. That competition was for the Presidential Physical Fitness Award, which is part of the American Alliance for Health, Physical Education, and Recreation's (AAHPER) Youth Fitness Program's motivational plan to increase the physical fitness levels of 9 to 17 year old American children.

In 1966, President Lyndon B. Johnson established the Presidential Physical Fitness Award. Added to the previously established Achievement Award Certificate and the Embossed Gold Merit Seal, the Presidential Physical Fitness Award provided strong incentives for Physical Education Programs to add the AAHPER Youth Fitness Tests to their curricula. To win the Achievement Award Certificate, the Embossed Gold Merit Seal or the Presidential Physical Fitness Award, children had to score above the 50th, 80th, or the 85th percentiles, respectively, on all of the tests.

Two reports by Hans Kraus and Ruth Hirshland created the controversy from which the AAHPER Youth Fitness Program evolved (Kraus and Hirshland, 1953 and 1954). In their reports they questioned the physical fitness levels of American children. Using the Kraus-Weber battery of six tests, Kraus and Hirshland measured: (1) children's range of motion during straight and flexed leg sit-ups; (2) the time that children sustained a supine double leg raise, a prone double leg raise, and a prone upper trunk raise; and (3) while maintaining a straight leg toe touch position, the distance above or below the platform on which the children stood that children could reach. The Low Back Clinic at Columbia Presbyterian Hospital developed the Kraus-Weber battery of tests. Comparing American children with Austrian, Italian and Swiss children, Kraus and Hirshland concluded that the American children were not as physically fit as foreign children.

Alerted by the Kraus-Hirshland studies, in 1956 President Dwight D. Eisenhower summoned a national conference to discuss the status of American children's physical fitness. President Eisenhower's conference inspired the 1956 AAHPER national convention to discuss what physical education teachers might do to upgrade the physical fitness of American children. As its first step, the AAHPER created the AAHPER Youth Fitness Project. As its second step, the AAHPER appointed a special AAHPER Research Council to direct the AAHPER Youth Fitness Project.

As their first directive, the AAHPER Research Council decided to conduct their own measurement of the physical fitness of American children. To select their battery of physical fitness tests, the AAHPER Research Council proposed the following five criteria: (1) the tests should be reasonably familiar to the students; (2) the tests should require little or no equipment; (3) the tests should apply equally to boys and girls; (4) the tests should apply to children aged 9 through 17; and, (5) the tests must measure different components of physical fitness. After considerable discussion and preliminary testing, the AAHPER Research Council chose the following physical fitness tests: (1) the normal and modified pull-up tests; (2) the straight leg sit-up test; (3) the shuttle run test;

(4) the standing long jum test; (5) the 50 yard dash; (6) the 600 yard run-walk; and (7) the softball throw for distance.

Directed by Professor Paul A. Hunsicker, in 1957 the University of Michigan's Survey Research Center coordinated the AAHPER Youth Fitness Program's first data collection. Eight thousand five hundred children participated in the 1957 sampling.

When the AAHPER Research Council compared those 1957 norms with the norms of other countries, they confirmed the Kraus-Hirshland finding that American children were not as physically fit as foreign children. Therefore, the AAHPER recommended that Physical Education Departments across the United States add the AAHPER Youth Fitness Tests to their curriculum.

To determine whether or not the AAHPER Youth Fitness Tests had effected the physical fitness of American children, in 1963 the AAHPER commissioned its second national survey. Directed again by Professor Paul A. Hunsicker, the University of Michigan's Survey Research Center coordinated the measurement of a different sample of 9,200 children. In the battery of tests for the second sampling, the AAHPER substituted the timed flexed arm hang for the girl's modified pull-up. When the AAHPER compared the norms from their 1965 and 1957 surveys, with the single exception of the 17 year old girl's softball throw for distance, for every age and every test, they found that both the boys and girls had significantly improved their scores.

In 1975, the AAHPER commissioned its third national survey. Directed jointly by Professors Paul A. Hunsicker and Guy A. Reiff, the University of Michigan's Survey Research Center coordinated the data collection. In their 1975 test battery, the AAHPER Research Council made three changes. They eliminated the softball throw for distance, substituted the flexed leg sit-up for the straight leg sit-up, and, in place of the 600 yard run-walk, for children aged 12 and under, they provided the options of a one mile or nine minute run, and for children aged 13 years and over, they provided the options of the one and one-half mile to 12 minute run. Also, as their sixth Youth Fitness Test criteria, the AAHPER Research Council required that the children should be able to assess their own performances.

When the AAHPER compared their 1975 and 1965 norms, they found that for both sexes across the eight chronological age classifications and the six tests, the children improved in only six of the 96 cells and they regressed in two cells. From examining those results, researchers might conclude that the AAHPER's Youth Fitness competition failed to stimulate American children to improve their physical fitness.

When comparing the norms of American children with those of foreign children or with those of American children from different data collections, because random sampling negated the effect of the variability in the bioloigical maturation, the AAHPER could classify the children by the same procedures the others had used. However, when the AAHPER established its Youth Fitness competition, the situation changed. Because the children now competed individually against each other, random sampling no longer negated the effect of the variability in the biological maturation. Therefore, the procedure by which the AAHPER classified the competitive match-ups should have eliminated the variability in biological maturation. Research evidence suggests that if at the start of its competition, AAHPER had equated the children on the basis of biological maturation, then they would have increased the likelihood of rewarding the quality of the children's efforts.

In 1966, the AAHPER used the following two procedures to classify the children for their Youth Fitness competition: (1) chronological age; and (2) the Nielson-Cozens Index. The Nielson-Cozens Index combined chronological age, height and weight. In 1975, because of a belief that height and weight added little to the equality of classifying children for motor proficiency norms, the AAHPER eliminated the Nielson-Cozens Index.

Need for the Study

There is need for a study to determine the proportion of the variability in adolescent males' biological maturation that is explained by increments in their chronological age, height, weight and a Sexual Maturation Indicator Value (SMIV) estimate. More specifically, this investigation examined the following questions;

Question 1: What proportion of the variance in the skeletal age of adolescent males is explained by increments in their chronological age, height and weight?

Question 2: Does an estimate of sexual maturation (SMIV) explain a significant additional proportion of the variance in adolescent males' skeletal age?

Question 3: Can boys' physical education teachers reliably assess the sexual maturation of adolescent males?

Question 4: How does an estimate of skeletal age (SA) that encompasses chronological age, height, weight and SMIV compare with chronological age when classifying adolescent males for motor proficiency norms?

Scope of the Study

To examine questions 1, 2 and 3, the author randomly selected 30 East Lansing, Michigan sixth, seventh, eighth and ninth grade male students to represent a rectangular distribution of sexual maturation. And, to examine question 4, every six months for three years the boys' physical education teachers assessed selected physical and motor characteristics of the entire male student body of East Lansing Junior High School and Mason Junior High School. Therefore, the scope of this study extends to sixth, seventh, eighth and ninth grade boys.

Limitations of the Study

The following factors may limit the interpretations of this study:

(a) Because this study included only sixth, seventh, eighth and ninth grade males, the findings apply only to males within the age range of 11 to 17 years.

(b) Although the AAHPER Youth Fitness Test Manual provided specific instructions for the administration of its battery of tests, the ability of the physical education teachers to conduct proper testing methodology may limit the interpretations.

(c) Whenever researchers measure children's motor proficiencies, a variety of environmental influences may limit interpretations. Some of those environmental influences are:

(1) the time of the day in which the children perform,

(2) the order in which the children perform,

(3) the presence of other children during the testing,

(4) the friction coefficient of the surface on which the children perform, and

(5) the temperature and humidity of the testing site.

(d) The physical education teachers received no special training. Presumably, they had no special competencies. Therefore, their ability to observe the (SMIV) checklist variables in a scientific manner may limit the interpretations.

(e) Except for the author's gratitude, the learning experience, and the advancement of their knowledge about the relationship between sexual maturation and motor proficiencies, those teachers received no reward for working in this study. Therefore, some might argue that the use of volunteer personnel limited the interpretations.

CHAPTER II

REVIEW OF LITERATURE

For the purpose of classifying adolescent males for athletic competitions, the author questions the use of chronological age. Because adolescent males of the same chronological age may be years apart in biological maturation, athletic competitions based on chronological age potentially reward boys on the basis of advanced biological maturation and not on the basis of effort or superior athletic ability. While this dissertation examines only the American Alliance of Health, Physical Education, and Recreation (AAHPER) Youth Fitness competition, the implications of the results extend to all athletic competitions for adolescent males that use chronological age to classify the boys.

One way that specialists in Child Growth and Development predict biological maturation of adolescent males is by estimating skeletal maturation from hand-wrist x-rays. In this study, Health, Physical Education, and Recreation Child Growth and Development Professor Vern Seefeldt used the Greulich-Pyle Radiographic Atlas of Skeletal Development of the Hand and Wrist to assess the skeletal age of 30 adolescent boys. Therefore, the first portion of the review of literature covers the history of the assessment of hand and wrist development.

Another way that specialists in Child Growth and Development predict biological maturation in adolescent males is by estimating the boys' sexual maturation. For this study, the author created a sexual maturation checklist that the male physical education teachers from East Lansing and Mason Junior High Schools used to assess the sexual maturation of their students. Therefore, the second portion of the review of literature covers the history of the assessment of adolescent males' sexual development.

The purpose of this study is to examine the procedures by which sponsors of athletic competitions classify adolescent males. The AAHPER's use of chronological age to classify adolescent males for motor proficiency norms is the focus of this investigation. Therefore, the third portion of the review of literature covers the relationship between the motor proficiency performances of adolescent males and various classification indices.

History of the Assessment of Hand and Wrist Skeletal Development

William K. Roentgen's December 1895 discovery of x-rays enabled researchers to study the progressive changes in the human skeleton. Prior to that discovery, knowledge of human skeletal development resulted from anatomists dissecting cadavers. Because 1) they received very few child cadavers and 2) they could not study their subjects longitudinally, anatomists provided limited information on skeletal development. However, by using x-rays, researchers could readily follow the skeletal development in children.

Four months after Roentgen's finding, on April 2, 1896 Professor Sydney Rowland of Germany x-rayed clinical patients, and, when describing the appearance of the bones of the children's hands and wrists, he recognized that those bones indicated the children's level of skeletal maturation (Pyle, Waterhouse, and Greulich, 1971). Later in 1896, another German researcher, Joseph Ranke, reported that x-rays presented a new method for determining children's biological age (Flory, 1936).

In the United States, Professor J.W. Pryor of the State College of Kentucky initiated the research on the x-ray assessment of hand and wrist skeletal development (Pryor, 1905). In 1905, from 70 x-rays of children aged 44 days to 21 years, Professor Pryor outlined the appearance order of the eight carpal ossification centers, and the 21 metacarpal, phalangial, ulnar, and radial epiphysial ossification centers. For boys and girls separately, Pryor summarized the ages at which their ossification centers appeared. He claimed that skeletal maturation occurred earlier in girls' hands than in boys' hands. This female maturation advantage was by days in early childhood, by months in early adolescence, and by years at the completion of skeletal development. Also, Pryor claimed that first born children's hands skeletally matured earlier than non-first born children's hands.

The following year, Professor Pryor compared his findings with the information on skeletal development contained in Cunningham's <u>Text Book of Anatomy</u>, Morris' <u>Human Anatomy</u>, Holden's <u>Human Osteology</u>, Gray's <u>Anatomy</u> and Poland's <u>Skiagraphic Atlas</u> (Pryor, 1906). He concluded that the skeletal ossification process of children's hands started earlier than reported in those references and that the appearance order of the children's epiphysial ossification centers differed from the appearance order reported in those references. Further, he emphasized that those references did not separate the data on boys from the data of girls.

Then, by using x-rays of identical twin girls and boy-girl twins, Pryor reaffirmed his claim that girls' hands skeletally matured earlier than boys' hands. Additionally, he emphasized that the distal ulnar and radial epiphyses developed simultaneously with the carpals, metacarpals, and phalanges, and therefore, he now included those sites with the skeletal maturation of the hand and wrist.

Pryor's next research topic effort delved into the hereditary nature of variations in carpal development (Pryor, 1907). After viewing 360 x-rays of children from 225 families, Professor Pryor observed that variation patterns in the appearance order of the carpals followed family lines. By recording the appearance order of the carpals for children from ten families, he showed that when a child's carpal appearance order differed from the order typical in the population, that child's brothers and sisters varied similarly. Additionally, Pryor noted the bilateral symmetry of the hand-wrist skeletal ossification process.

After expanding his sample of children aged 14 years and under to 266 girls and 288 boys, Pryor reaffirmed that the skeletal ossification process began earlier than previously reported (Table

2.1) and, that the carpal appearance order differed from those orders given in previous reports (Table 2.2), (Pryor, 1908). From the information in Table 2.1 he hypothesized that the children's carpal appearance order to be: (1) Os Magnum; (2) Unciform; (3) Cuneiform; (4) Semilunar; (5) Scaphoid; (6) Trapezoid; (7) Trapezium; and (8) Pisiform. Additionally, Pryor reported the sex, age, order of birth, carpal appearance order, and millimeter diameter size of each carpal for 136 girls and 153 boys.

Table 2.1

The Age by Sex of Carpal Appearance in Children (Pryor, 1908)

Carpal	Sex	Age of Appearance
Os Magnum	Female	Between 3-4 months
(Capitate)	Male	Between 7-8 months
Unciform	Female	Between 3-4 months
(Hamate)	Male	Between 7-8 months
Cuneiform	Female	13th or 14th month
(Triquetral)	Male	early in the 3rd year
Semilunar	Female	Late in the 3rd year or early in the 4th year
(Lunate)	Male	Between the 4th and 5th years
Scaphoid	Female	About 4 years
	Male	About 6 years
Trapezoid	Female	Between the 4th and 5th year
	Male	Between the 5th and 6th year
Trapezium	Female	Between thd 4th and 5th year
	Male	Early in the 8th year
Pisiform	Female	Between the 9th and 10th year
	Male	Between the 12th and 13th year

In a 1925 report, Professor Pryor examined x-rays of 64 boys aged 150 to 274 months and 81 girls aged 144 to 260 months. In table form, Pryor recorded the ages, sex, and an epiphysial rating of whether he rated Distinct or Indistinct for all 145 children.

Table 2.2

The Frequency of the Carpal Appearance
Order in 243 Children (Pryor, 1908)

CARPALS	1ST	2ND	3RD	4TH	5TH	6TH	7TH	8TH	TOTAL
Os Magnum	238	5	0	0	0	0	0	0	243
Unciform	5	236	0	0	0	0	0	0	241
Cuneiform	0	0	219	10	0	0	0	0	229
Semilunar	0	0	12	176	8	2	6	0	204
Scaphoid	0	0	0	3	80	27	19	0	129
Trapezoid	0	0	0	8	36	59	21	0	124
Trapezium	0	0	0	7	30	31	51	0	119
Pisiform	0	0	0	0	0	0	0	19	19

Pryor noted the ages at which the epiphysial centers appeared (Table 2.3). Responding to criticism that he had not compared actual bones with x-rays of those same bones, Pryor countered that x-rays showed the bones' texture and the very faint epiphysial lines more clearly than the naked eye could detect (Pryor, 1925). In a 1928 report, Professor Pryor presented an historical review of his 30 years of research into the skeleton's ossification process.

Unaware of Pryor's reports, Boston Physician Thomas Morgan Rotch pondered an anatomical age (Rotch, 1908). X-ray technician A.W. George of Boston's Children's Hospital had alerted Dr. Rotch to the anatomic age concept. From examining over 500 x-rays of children's hands, wrists, elbows, shoulders, hips, knees, and ankles, Dr. George decided that anatomic age indicated the biological age differences between children better than chronologic age. And he concluded that the carpal bones indicated anatomic age better than any other bones. Working from 200 x-rays of children's carpal bones, Dr. Rotch developed his Anatomic Stage Index. Because he wanted

Table 2.3

The ages by sex that the hand-wrist epiphysial centers appeared (Pryor, 1925)

Epiphysial Center	Sex	Time of Appearance
1st Metacarpal	Female	Early in the 3rd year
	Male	Early in the 3rd year
2nd Metacarpal	Female	Late in the 1st year
	Male	Early in the 2nd year
3rd Metacarpal	Female	Early in the 2nd year
	Male	Late in the 2nd year
4th Metacarpal	Female	Early in the 2nd year
	Male	Late in the 2nd year
5th Metacarpal	Female	Early in the 2nd year
	Male	Late in the 2nd year
Middle Phalange:		
First Row	Female	Late in the 1st year
	Male	Early in the 2nd year
Second Row	Female	Early in the 2nd year
	Male	Late in the 2nd year
Third Row	Female	Late in the 2nd year
	Male	Late in the 3rd year
Ring Phalange:		
First Row	Female	Early in the 2nd year
	Male	Late in the 2nd year
Second Row	Female	Early in the 2nd year
	Male	Early in the 3rd year
Third Row	Female	Late in the 2nd year
	Male	Late in the 3rd year
Index Phalange:		
First Row	Female	Early in the 2nd year
	Male	Late in the 2nd year
Second Row	Female	Early in the 2nd year
	Male	Early in the 3rd year
Third Row	Female	Late in the 2nd year
	Male	Late in the 3rd year
Little Phalange:		
First Row	Female	Early in the 2nd year
	Male	Late in the 2nd year
Second Row	Female	Late in the 2nd year
	Male	Early in the 3rd year
Third Row	Female	Late in the 2nd year
	Male	Late in the 3rd year
Thumb Phalange:		
First Row	Female	Late in the 2nd year
	Male	Early in the 3rd year
Second Row	Female	Early in the 2nd year
	Male	Late in the 2nd year
Radius	Female	About the 8th month
	Male	About the 15th month
Ulna	Female	Between the 6th and 7th year
	Male	Between the 7th and 8th year

to separate anatomic from chronologic age, Rotch alphabetized his Index (Table 2.4).

Table 2.4

Alphabetized Anatomic Stage Index
(Rotch, 1908)

Chronologic Age	Anatomic Stage	Description
6 months	A	Os Magnum and Unciform present
33 months	B	Distal Radial Epiphysis appears
33 months	C	Cuneiform appears
27 months	D	Semilunar appears
42 months	E	Either the Trapezium or the Scaphoid appears
66 months	F	Both the Trapezium and the Scaphoid appear
78 months	G	Trapezoid appears
72 months	H	Distal Ulnar epiphysis appears
81 months	I	Total carpal development advances
99 months	J	Total carpal development advances
135 months	K	Pisiform appears
153 months	L	Pisiform is very clear and total carpal development advances
162 months	M	Pisiform is almost as large as the Cuneiform

He hypothesized that when classifying children for school, athletics, or child-labor, anatomic age defined children's readiness better than chronologic age. To support his claim, Rotch cited Dr. C. Ward Crampton's work on the superiority of physiologic age over chronologic age (Crampton, 1908a). Because researchers could use anatomic age from infancy to adulthood, Rotch contended that anatomic age estimated biological age better than any other method.

After a series of communications with Professor Pryor, in his next paper, Rotch deferred to Pryor's anatomic expertise and corrected two errors in his first report (Rotch, 1909). He reversed the appearance order of the Trapezium and the Trapezoid, separated the ossification processes of boys and girls, and included Pryor's Anatomic Stage Index in his paper (Table 2.5). Rotch said that boys and girls alike could use his Anatomic Stages, but to determine the chronological age that corresponded with each Anatomic Stage, he had to separate boys from girls. In a huge table containing data from 129 boys and girls aged one month to 13 years rated as intellectually bright, Rotch provided their sex, height, weight, number of teeth, and Anatomic Stage. From that table, he concluded that chronological age, height, weight, number of teeth, and parental or guardian statements insufficiently attested to the children's readiness for school, athletics, and physical work.

Table 2.5

Pryor's Anatomic Stage Index (Rotch, 1909)

Sex	Chronologic Age (months)	Anatomic Stage	Description
Female	6	A	Magnum and Unciform present
Male	12	A	
Female	18	B	Distal radial epiphysis
Male	24	B	appears
Female	30	C	Cuneiform appears
Male	30	C	
Female	42	D	Semilunar appears
Male	51	D	
Female	51	E	1st of the Scaphoid, Trapezoid,
Male	63	E	or Trapezium appears
Female	57	F	2nd of the Scaphoid, Trapezoid,
Male	69	F	or Trapezium appears
Female	63	G	3rd of the Scaphoid, Trapezoid,
Male	81	G	and Trapezium appears
Female	75	H	Distal Ulnar epiphysis appears
Male	87	H	
Female	90	I	Everything advances
Male	108	I	
Female	102	J	Everything advances
Male	120	J	
Female	123	K	Pisiform appears
Male	153	K	

Although applauding Crampton's work on physiologic age, Rotch reiterated that because Crampton's Index required someone to view the boys' public hair, the schools and courts could not use his procedure. Also, he said that because Crampton's Index applied only to the adolescent aged boys, anatomical age estimated biological maturation better than Crampton's Index. Nevertheless, Rotch credited Crampton for disproving chronologic age as a classification procedure for children's readiness for school, athletics, and child labor. Rotch again called for child labor laws based on anatomic age and he advised educators to consider epiphysial development as an educability indicator. To accentuate that point, Rotch suggested that when educators found children whose mental development exceeded their epiphysial development, to enable epiphysial development to catch up, educators should withdraw those children from mental exercises.

In May of 1909, the Secretary of the Navy asked Dr. Rotch to prepare an Anatomic Index that Naval officers could use to determine the anatomic age of Anapolis Cadets. From x-rays of Naval Cadets aged 16 to 21 years old, Drs. Rotch and Harold Wellington Smith concluded that after boys reached 13 years of age, researchers should use the boys' epiphysical unions to describe their Anatomic Age. They found that the epiphyses that appeared first also ossified first (Rotch and Smith, 1910). Using the distal radial epiphysis as their primary criteria, they proposed the following four levels of epiphysial union: (1) distal radial epiphysial union had just begun; (2) distal radial epiphysis had ossified one-quarter across the radial shaft; (3) one-half distal radial epiphysis has ossified; and, (4) distinct distal radial epiphysial lines.

Later that year, Harvard Pediatrics Professor T.M. Rotch used his Anatomic Index to correlate biological age with school grade level (Rotch, 1910). Rotch detailed the Anatomic Stage that he believed the children in each grade should have attained (Table 2.6). To demonstrate the concept, he provided hand-wrist x-rays of three boys aged 7 to 10 years old. Because those boys' Anatomic Stages were identical, he stated that they belong in the same school grade.

Table 2.6

Academic Grade-Anatomic Stage (Rotch, 1910)

Academic Grade	Chronologic Age	Anatomic Stage
K1	4 years	E
K2	5 years	F
1st	6 years	G
2nd	7 years	G+ and H
3rd	8 years	H+ and I
4th	9 years	I+ and J
5th	10 years	J+
6th	11 years	K
7th	12 years	L
8th	13 years	M

Under the auspices of the Committee on Physical Welfare of School Children, New York City Physician C. Ward Crampton studied the anatomic and physiologic ages of males (Crampton, 1910). He said that Rotch's Anatomic Index was too irregular and subjective, and, therefore, not reliable. Also, Crampton disagreed that the skeletal development of the boys' hands indicated their total skeleton's development.

Nevertheless, to test Rotch's Anatomic Stages, Dr. Crampton gave permission and assistance to New York physicians Eli Long and E.W. Caldwell to analyze hand-wrist x-rays of 200 New York City Department of Education children aged 13 to 173 months (Long and Caldwell, 1911). Although they did not compare carpal ossification with total skeletal ossification, because in one-third of their sample the right hand x-ray differed from that of the left hand, Long and Caldwell doubted that carpal ossification accurately reflected total skeletal development. Because Rotch's Anatomic Stages I and J allowed x-ray readers to subjectively evaluate the relationship between the children's increasing bone calcification and their decreasing carpal interspace area, Long and Caldwell condemned Rotch's Index. Consequently, they graded their x-rays in numerical order from least to greatest skeletal development. Long and Caldwell concluded that although Rotch's Anatomic Stages unsatisfactorily quantified skeletal development, they hoped that the skeleton's ossification centers would eventuate such an index.

Continuing the work he and Dr. Rotch started in 1909, Boston physician Harold Wellington Smith x-rayed 914 Naval midshipmen aged 17 to 23 years old (Smith, 1913). Using the 21 epiphyses of the hand and wrist, Smith outlined their appearance order and age of completed union. Within the epiphysial ossification process, Smith differentiated eight stages (Table 2.7). For each hand-wrist epiphysis, Smith plotted curves of the ages of the subjects at each of his eight stages. Smith expected those curves to define a uniform epiphysial development sequence. When those epiphysial development

Table 2.7

The Descriptions of the Eight Epiphyseal Development Stages in Adolescent Males (Smith, 1913)

Stage	Descriptions
1	No sign of beginning epiphysial union
2	A few imbrications have appeared in the clear epiphysial cartilage
3	The imbrications have interrupted the clear epiphysial cartilage and have reached the proximal side
4	Except for the periphery, the imbrications have fused into interlacing multiple lines that occupy the cartilage space
5	The interlacing multiple lines have contracted into an irregular thick line with deep peripheral notches
6	The thick line has become a distinct, narrow line that extends the full width of the bone except for a slight peripheral notch
7	Only portions of the line and traces of the peripheral notch remains
8	The bony matrix of the shaft is contiguous with the epiphysial end

curves did not define a uniform sequence. Based on those results, he concluded that his stages did not determine epiphysial development correctly.

For the first 25 years after Roentgen's discovery of x-rays, individual researchers such as Professor J.W. Pryor and Dr. T.M. Rotch did the majority of the research on children's hand and wrist skeletal development. But, beginning around 1920, University research centers became involved in the assessment of skeletal maturation. While there were some one-time studies by individuals thereafter, from 1920 on, children research centers such as those at the University of Iowa, Harvard University, the University of Chicago, and Cleveland's Western Reserve University dominated the study of hand and wrist skeletal development.

Directed by University of Wisconsin Professor C.R. Bardeen, University of Wisconsin Graduate Student Earnest Donald requested Roentgenology Instructor Dr. Howard Curl to x-ray 96 boys and 71 girls aged 42 to 138 months (Bardeen, 1921). Donald correlated the children's carpal ossification with their chronological age, height, and weight. After comparing Donald's results with Professor Pryor's results, Bardeen concluded that the age differences in Donald's sample were not as large as the differences in Pryor's sample. To make finer subdivisions of carpal ossification than Pryor's Distinct or Indistinct categories, Bardeen created four carpal ossification classes (Table 2.8). Using those four ossification classes as his

Table 2.8

Four carpal ossification classes (Bardeen, 1921)

Class	Description
A	Large, dense carpal ossification center
B	Not as developed as Group A
C	More developed than Group B
D	Barely visible carpal ossification center

base, Bardeen divided the carpal ossification process into 11 stages (Table 2.9). In only four of 96 boys and two of 71 girls did Bardeen find significant differences in right and left hand ossification stage. Because when he compared his sample's age of carpal center appearances with height and weight, he found no relationship between those variables, Bardeen said that based on present knowledge, school officials should not delay children's education because their carpal centers appeared late.

Table 2.9

Eleven Carpal Ossification Stages
(Bardeen, 1921)

Stage	Description
1	Class C Capitate and Hamate
2	Class C Capitate, Hamate, and Triquetral
3	Class C Capitate, Hamate, Triquetral, and Lunate
4	Either Trapezium, Trapezoid, or Scaphoid appeared
5	Class D Trapezium or absent. Pisiform absent. All other carpals class B or C.
6	Class C Trapezium and Trapezoid. Pisiform absent. All other carpals class B.
7	Class D Pisiform or absent. All other carpals class B.
8	Class D Pisiform or absent. Class A Capitate and Hamate. All other carpals class B.
9	Class D Pisiform. All other carpals class A.
10	Class B Pisiform. All other carpals class A.
11	All carpals class A.

In his three year study of 31 girls aged 10 to 13 years and 36 boys aged 11 to 13 years from the University of Iowa Junior High Schools, University of Iowa Professor Bird T. Baldwin analyzed the anatomical growth of children's carpal bones (Baldwin, 1921). Baldwin searched for a completely objective hand-wrist x-ray evaluative method. First, he tried measuring the perimeters of the carpals. Next, he tried tracing the outlines of the carpals on millimeter graph paper. Using a planimeter, Baldwin eventually decided to measure the total exposed area of the carpals. Because the Pisiform was not present in all of his subjects, Baldwin excluded the Pisiform from his measurement of the total exposed carpal area.

Using Pearsonian correlation coefficients to compare total carpal area with height and weight, Baldwin found that the boys' correlations were .879 and .755 respectively, and the girls' correlations were .729 and .766 respectively. Although he found that the carpal bones in the right and left hands of some children had great individual differences, Baldwin concluded that on the average the carpal areas for both hands were equivalent. For boys and girls of the same chronological age, Baldwin found that the girls' carpal area measurements exceeded the boys' carpal area measurements.

Following Professor Pryor's example and because he had non-identical twin sons, Baldwin studied four sets of twins. Baldwin concurred with Pryor that in boy and girl twins, the girls' hands matured earlier than the boys' hands. Additionally, Baldwin observed that children's hand-wrist x-rays showed that their carpals had distinct stages of growth. Therefore, Baldwin promised that he would study the 280 University of Iowa laboratory school children to standardize their carpal growth stages into male and female anatomical calendars.

The next year Professor Baldwin reported on the mental growth curves of normal and superior children (Baldwin, 1922). Among the measures that he took on his subjects, Baldwin assessed an x-ray or anatomic stage. Again, he determined the subjects' anatomic stage by planimetering the total area of the carpal bones. Baldwin correlated chronological age, weight, height, mental age, and total exposed surface area of their carpal bones (Table 2.10). Baldwin claimed that carpal age related closely to physiologic changes.

Frances Lowell and Herbert Woodrow examined hand-wrist x-rays of 402 Minneapolis and St. Paul children aged 66 to 138

Table 2.10

Correlation Table for Chronological Age, Weight, Height, Mental Age, and Carpal Age (Baldwin, 1922)

	Chronologic Age	Weight	Height	Mental Age	Carpal Age
Chronologic Age	1.00				
Weight	.84	1.00			
Height	.98	.86	1.00		
Mental Age	.88	.71	.89	1.00	
Carpal area	.92	.88	.92	.83	1.00

months (Lowell and Woodrow, 1922). After separating the x-rays by age and sex, they ranked the x-rays from best developed to least developed. For that ranking, they considered the surface area of the carpals, the shape of the carpals, the number of carpals, the identity of the carpals present, and the carpal area relative to the total wrist area. They selected the median x-ray as the standards. Because this technique ultimately provided today's hand-wrist skeletal development standards, researchers should recognize Lowell and Woodrow as the pioneers of the concept. Lowell and Woodrow described those median x-rays (Table 2.11).

In addition to those carpal measurements, Lowell and Woodrow determined the number of permanent teeth. When they correlated the carpal development with number of permanent teeth, they found a .20 correlation. Similarly, when they correlated the carpal development with school grade, they found another low correlation. Because of the low carpal development-school grade correlation, they concluded that they could not agree with Rotch's concept of using anatomic age

Table 2.11

X-Ray Standard Descriptions (Lowell and Woodrow, 1922)

Age (Years)	Description
Boys: 5½	5 carpals present, one very small. No ulnar epiphysis. Carpal area = 2.9 square centimeters.
Girls:	7 carpals. No ulnar epiphysis. Carpal area = 3.4 sq. cms.
Boys: 6½	7 carpals, two very small. No ulnar epiphysis. Carpal area = 3.1 sq. cms.
Girls:	7 fair size carpals. Ulnar epiphysis extends one-quarter the end of the ulna. Carpal area = 416 sq. cms.
Boys: 7½	Ulnar epiphysis extends one-sixth the end of the ulna. Carpal area = 5.7 sq. cms.
Girls:	Ulnar epiphysis extends sixty percent the end of the ulna. Carpal area = 5.9 sq. cms.
Boys: 8½	Ulnar epiphysis extends sixty percent the end of the ulna. Carpal area = 5.9 sq. cms.
Girls:	Ulnar epiphysial-diaphysial contact. Trapezium-Trapezoid overlaps. Carpal area = 7.8 sq. cms.
Boys: 9½	Oval ulnar epiphysis. Little or no space between the Os Magnum and the Unciform. Carpal area = 6.1 sq. cms.
Girls:	Radial epiphysis contacts shaft. Ulnar styloid begins. Ulnar epiphysis better developed. Carpal area = 9.7 sq. cms.
Boys: 10½	Better ulnar epiphysial development. Better carpal development. Carpal area = 8.9 sq. cms.
Girls:	Well-developed ulnar styloid. Small pisiform. Carpal area = 9.7 sq. cms.
Boys: 11½	Ulnar styloid begins. Tiny pisiform. Carpal area = 9.6 sq. cms.
Girls:	Good-size Pisiform. Good Carpal compactness. Carpal area = 9.6 sq. cms.

to place children in school grades, nor could they agree with Crampton's use of physiological age for the same purpose.

Under a Commonwealth Fund grant, Harvard University Professor Walter F. Dearborn of the Harvard Graduate School of Education examined the mental and physical development of over 5,000 children. With Professor Dearborn's permission, Professor Daniel Alfred Prescott of Harvard University's Psycho-Educational Clinic analyzed

3,000 hand-wrist x-rays (Prescott, 1923). Professor Prescott wanted to develop an objective method for assessing children's hand and wrist ossification stages, and using that objective method, he wanted to establish objective hand-wrist skeletal development norms.

Because he believed that Pryor and Rotch did not discriminate skeletal development increments in an objective manner, Prescott rejected their indices. Although Prescott credited Baldwin with the first attempt at establishing objective procedures, because Baldwin failed to adjust his total carpal area index for the size of the children's hands, he concluded that Baldwin's carpal area index insufficiently estimated skeletal age. Prescott also praised Lowell's and Woodrow's pioneer work, but he stated that because their carpal development also failed to account for the variable sizes of children's hands, Lowell's and Woodrow's procedure also insufficiently estimated skeletal age.

For his initial experimental sample, Prescott selected 38 boys and girls aged 83 months to 114 months. He and an unidentified co-worker independently ranked those 38 x-rays from least to most developed. Because their rankings correlated at .98+, Prescott contended that they had ranked those x-rays correctly. He found bilateral symmetry in over 80 percent of the hands. Because in this initial sample Prescott found that girls matured earlier than boys, he repeated the experiment with a second sample of 39 boys aged 74 to 76 months. Even though he had eliminated the sex and age variabilities, Prescott and his unidentified co-worker again ranked the x-rays similarly. Therefore, Prescott believed that any objective measurement that ranked the x-rays in the same order as their inspectional ranking order would be the objective measurement method.

To find his objective method, the factors that Prescott considered were: (1) the number, size, and shape of the carpals; (2) the quality of the carpal ossification; (3) the articulation between the carpals; (4) the unossified area in the wrist, (5) the size of the hand; (6) the number of the metacarpal, phalangial, ulnar, and radial epiphysis present; and (7) the extent of their epiphysial-diaphysial union. Prescott believed that only when he accounted for hand size could ossified carpal area demonstrate the extent of ossification. Because planimetering the carpals directly on the x-ray proved unreliable, Prescott projected the x-ray onto drawing paper. When the projected image measured five times the x-ray original, he traced the carpal bones.

In his search for the objective measurement method, Prescott drew nine lines on each x-ray's tracing (Table 2.12).

In the beginning, Prescott used only Line A (the trapezium-hamate diameter), Line B (the trapezium-triquetral diameter), and Line C (the scaphoid-hamate diameter). When he found that those three measurements adequately separated skeletal developmental levels, he combined those measures into ratios. However, he found that those ratios also lacked discriminative ability. At this point Prescott realized that he had to account for the varying hand sizes. Therefore, by connecting the following four points, Prescott created the last five of his lines; (1) the medial-most point on the proximal end of the shaft of the first metacarpal, (2) the medial-most point

Table 2.12

The Nine Measurement Lines that Prescott
Drew on the Hand-Wrist X-Rays
(Prescott, 1923)

Line	Description
A	The widest diameter from the outside of the trapezium to the outside of the hamate
B	The widest diameter from the outside of the trapezium to the outside of the triquetral
C	The widest diameter from the outside of the scaphoid to the outside of the hamate
D	The distance between the first metacarpal's landmark and the fifth metacarpal's landmark
E	The distance between the fifth metacarpal's landmark and the radius' landmark
F	The distance between the first metacarpal's landmark and the radius' landmark
G	The distance between the radius' landmark and the ulna's landmark
H	The distance between the ulna's landmark and the first metacarpal's landmark
K	The widest diameter between the hamate and the triquetral

on the proximal end of the shaft of the fifth metacarpal, (3) the lateral-most point on the distal end of the shaft of the radius, and (4) the medial-most point on the distal end of the shaft of the ulna.

Theorizing that he needed a ratio of the yet-to-ossify to the already-ossified area, Prescott hypothesized that Line D times Line E represented the yet-to-ossify area and Line A times Line K represented the already-ossified area. Because 1) this DE/AK ratio decreased sytemmatically with increasing age from 5.10 down to 1.58 and 2) the DE/AK ratio correlated .85 with their inspectional ranking order, Prescott believed that he was on the right path to determining the objective measurement method. However, when he found that Baldwin's, and Lowell's and Woodrow's, carpal area measures correlated .93 with Prescott's inspectional ranking order, Prescott decided to search for a better ratio.

Consequently, Prescott added the following three measures to his x-ray tracings: (1) D = the sum of the widest diameters of all of the carpals; (2) W = the diameter of the wrist (Line D); and (3) inter-c = the sum of the shortest distances between the carpals. From those new criteria Prescott created three ratio combinations. When he correlated those new ratios with their inspectional ranking he found all of the correlations exceeded .94 (Table 2.13). In

Table 2.13

The Correlations that Prescott Found for His
Carpal Ossification Ratios
(Prescott, 1923)

Experimental Groups	D/W	inter-c/W	D/inter-c
38 boys and girls aged 88-114 months	.97	.98	.99-
39 boys aged 74-76 months	.96+	.94	.96-
46 boys aged 77-79 months	.95	.96	.97+
84 boys aged 86-88 months	.95	.95	.97

addition to his first two experimental groups, Prescott used two more experimental samples to correlate his findings. Therefore, Prescott contended that all three ratios provided a reliable, significant, and discriminating measure of carpal development. However, because D/W required only nine x-ray tracing measurements and less than three minutes per x-ray, whereas the inter-c and D/ inter-c ratios required five and eight minutes, respectively, Prescott adopted the D/W ratio as the objective measure of assessing anatomical age.

Using the D/W ratio, Prescott measured two more experimental groups to further test his objective measurements' ranking order with their inspectional ranking order. From the x-rays of 92 boys aged 83-86 months and 59 girls aged 83-86 months, he found that the D/W ratio correlated 98+ and 96+, respectively. Prescott called his D/W ratio values the Anatomical Index. By determining the Anatomic Index for the median x-ray of each of his age groups, Prescott indexed the anatomic ages. To create Table 2.14, the author interpreted the charts that Prescott provided. Prescott claimed that

Table 2.14

The number, sex, age range, and anatomical indices of Prescott's sample (Prescott, 1923)

Number of Subjects	Sex of Subjects	Age range of Subjects	Anatomic Index
114	Female	69-74 months	1.70
137	Female	75-80 months	1.82
135	Female	81-86 months	1.92
158	Female	87-92 months	2.03
86	Female	93-98 months	2.13
38	Female	108-119 months	2.50
37	Female	120-131 months	2.68
40	Female	132-143 months	2.90
26	Female	144-155 months	3.06
52	Female	156-167 months	3.14
46	Female	168-179 months	3.18
52	Female	180-191 months	3.18
28	Female	192-203 months	3.26
36	Female	204-215 months	3.18
15	Female	216-227 months	3.20
112	Male	69-74 months	1.23
162	Male	75-80 months	1.43
175	Male	81-86 months	1.54
149	Male	87-92 months	1.70
108	Male	93-98 months	1.84
17	Male	108-119 months	2.00
29	Male	120-131 months	2.38
24	Male	132-143 months	2.55
34	Male	144-155 months	2.69
36	Male	156-167 months	2.98
41	Male	168-179 months	3.10
50	Male	180-191 months	3.18
35	Male	192-203 months	3.10
28	Male	204-215 months	3.28
15	Male	216-227 months	3.24

the uniform distribution of the Anatomic Indices established the validity and merit of his D/W ratio.

Supervised by Western Reserve University Anthropology Professor T. Wingate Todd, Masters' Degree Candidate Paul H. Stevenson examined110 skeletons from persons who died at ages 15 to 28 years old (Stevenson, 1924). Stevenson wanted to determine the sequence and duration of epiphysial unions. To acquire a general understanding of epiphysial union, Stevenson developed four epiphysial union phases (Table 2.15). For the 28 epiphyses associated with the humerus, radius, ulna, femus, tibia, scapula, hip, ribs, and clavicle, Stevenson discussed the age of union of the diaphyses with their epiphyses. Considering the ulna and radius together, Stevenson noted that their proximal epiphyses united earlier than their distal epiphyses. He determined that the distal epiphyses of the radius and ulna united at age 19 years.

In 1921, the University of Chicago Department of Education's Laboratory Elementary School staff began x-raying the hands and wrists of their sample. After a year of collecting those x-rays, University of Chicago doctoral candidate Thomas M. Carter analyzed those x-rays. After two years of work, Professor Frank N. Freeman and Mr. Carter reported their carpal ossification ratio (Freeman and Carter, 1924). While they referenced the Rotch, Woodrow, and

Table 2.15

The Four Epiphyseal Union Phases that Stevenson Described (Stevenson, 1924)

Phase	Description
1 (No union)	The epiphysial and diaphysial margins project saw-tooth-like external margins. During macenation, the epiphysis may separate from its diaphysis.
2 (Beginning Union)	The epiphysial line develops and occasionally the two margins knit together.
3 (Recent union)	Active bone union completed, but, a fine reddish demarcation line continues. Most difficult phase to assess.
4 (Complete union)	Completed epiphysial union. Do not mistake the capsular attachment line with the epiphysial line.

Baldwin articles, they did not review Prescott's work. However, in a footnote, they said that they now had Prescott's article, but they had completed their work prior to receiving it.

Freeman and Carter believed that carpal development indicated the children's immediate position in the skeletal maturation process. Because they found carpal diameter measures too irregular, they decided that Baldwin's total carpal area measure might be more reliable. Because general skeletal size influenced total carpal area, they concluded that their measurement procedure had to account for the varying hand sizes. Therefore, they conceived a carpal ossification ratio.

In their carpal ossification ratio, their numerator factor was the ossified carpal area measure, and their denominator was a carpal quadrilateral area measure. They hypothesized that their carpal quadrilateral area measure adjusted for the general skeletal size. Although later in their article they said that they would specify the points of their carpal quadrilateral, their only description was that "points on the bones of the extremities of the arm and of the hand" formed the carpus quadrilateral.

After analyzing the hand-wrist x-rays of 20 girls and 20 boys at each age from five years through 17 years, Freeman and Carter provided separate tables for: (1) the carpal quadrilateral; (2) the ossified carpal area; and (3) the carpal ossification ratio. For the boys, the ossified carpal area ranged from 45 mm. at age five to 355 mm. at age 17. Therefore, their carpal ossification ratio ranged from 45/160 = 0.28 at age five to 355/330 = 1.07 at age 17.

Four months after Freeman and Carter reported their method, Arthur I. Gates reported on his 1922-23 experimental study at the Horace Mann School (Gates, 1924). Because Gates needed a measure of anatomical development, he reviewed the literature for an objective procedure. After reading Baldwin's work, Gates decided that Baldwin failed to account for varying hand sizes. Therefore, Gates hypothesized an ossified carpal area to a carpal quadrilateral area ratio.

To represent his carpal quadrilateral area, Gates selected points on the extreme outer corners of the first metacarpal, fifth metacarpal, ulna, and radius. By combining both hands and dividing the ossified areas, Gates computed his Ossification Percentage. For 30 boys and 28 girls with an average of 68 months and an average deviation of 12 months, Gates' Ossification Percentage ranged from 11 to 44 percent, with a mean of 25 percent. For 26 boys and 30 girls with an average age of 115 months and an average deviation of 10 months, Gates' Ossification Percentage ranged from 41 to 72 percent with a mean of 55 percent.

To show that when compared to the Freeman-Carter Ossification method any inspectional method inadequately estimated skeletal age, Joseph C. McElhannon reported on a reliability study that he participated in at the University of Chicago (McElhannon, 1926). For 20 boys and 20 girls at each age from seven years to ten years, McElhannon first inspected the x-rays and ranked them from least to most mature. McElhannon explained that the landmarks of the Freeman-Carter carpal quadrilateral were the "proximal points of the radius, ulna, fifth metacarpal, and the epiphysis proximum of the first metacarpal," and that the ossified carpal area included the Multangulum Majus, Multangulum Minus, Naviculare, Lunatum, Triquetrum, Capitatum, Hamatum, Pisiforme, and the distal radial and ulna epiphyses. In addition to measuring the area of the carpals and the carpal quadrilateral by the Freeman-Carter method, he observed the shape, number, and identity of the carpals present. Then, to determine which inspectional method ranked the x-rays the best, McElhannon used the following six different criteria; (1) the development of the Multangulum Majus, Multangulum Minus, and Novaculare; (2) the development of the Lunatum and Triquadrum; (3) the development of the Capitatum and Hamatum; (4) the development of the distal radial epiphyses and the distal ulnar epiphysis; (5) the fusion of the distal ulnar epiphysis; and (6) general observation.

Because a composite of all of those inspectional criteria correlated almost perfectly with his initial general inspectional ranking, McElhannon determined that all inspectional ranking methods were equivalent. Therefore, he concluded that having several inspectional criteria failed to increase the rater's ability to discern the skeletal maturation.

Next, he correlated his inspectional method with the Freeman-Carter method (Table 2.16). McElhannon concluded that as the carpal

Table 2.16

The Correlation Values Between McElhannon's Inspectional Method and the Freeman-Carter Method (McElhannon, 1926)

7 year boys	7 year girls	8 year boys	8 year girls	9 year boys	9 year girls	10 year girls	14 year boys
.85	.91	.82	.73	.75	.62	.41	.53

bones began to overlap and fill the carpal quadrilateral, the reliability of inspectional methods decreased. Therefore, because as carpals grew, the inspectional estimate decreased, McElhannon advised that workers reject inspectional methods and use a scientific measurement method (McElhannon, 1926).

Later that same year, Thomas M. Carter also reported on the reliability of the Freeman-Carter method (Carter, 1926). In this article, Carter described the following four points as the Freeman-Carter carpal quadrilateral landmarks: (1) the extreme outside point of the distal ulnar shaft; (2) the most proximal point of the fifth metacarpal; (3) the nearest point on the first metacarpal epiphysis; and (4) the extreme outside point on the distal radial shaft. To verify the Freeman-Carter measurement method, Carter and fellow graduate student E.J. Brown independently measured 15 children's x-rays by the Freeman-Carter procedure. Then they x-rayed those children a second time and again independently measured their x-rays. Their test-retest correlations were .97 and .96, respectively. Carter concluded that those correlations verified the reliability of the Freeman-Carter method.

Additionally, Carter reported that three other University of Chicago graduate students also tested the Freeman-Carter method. Using x-rays of 20 boys and 20 girls at each age of 9, 10 and 11 years old, W.H. Buchanan correlated an inspectional method with the Freeman-Carter method (Table 2.17). J.G. McElhannon and Everett

Table 2.17

The Correlation Values Between Buchanan's Inspectional Method and the Freeman-Carter Method (Carter, 1926)

	9 yr old boys	9 yr old girls	10 yr boys	10 yr girls	11 yr boys	11 yr girls
Buchanan's Correlations	.85	.90	.78	.61	.69	.58

Davis studied x-rays of 20 boys and 20 girls at each age of seven, eight and nine years, and 20 girls aged ten years and 20 boys aged 14 years (Table 2.18)

Table 2.18

The Correlation Values Between the Inspectional Methods of McElhannon and Davis and the Freeman-Carter Method (Carter, 1926)

	7 yr boys	7 yr girls	8 yr boys	8 yr girls	9 yr boys	9 yr girls	14 yr boys	10 yr girls
McElhannon's within-rater	.85	.91	.82	.73	.75	.62	.41	.53
Davis' within-rater	.86	.62	.83	.73	.78	.73	.87	.92
McElhannon-Davis between-raters	.49	.81	.91	.91	.88	.69	.51	.69

Table 2.17 shows why McElhannon concluded that the accuracy of his inspectional method decreased as the carpals filled the Freeman-Carter carpal quadrilateral. However, McElhannon's interpretation failed to hold for Davis. In Fact, Davis' two highest values were for the oldest two groups.

In 1927, University of Minnesota Professors Richard E. Scammon and Gordon H. Scott studied the accuracy of the 7-inch Amsler planimeter that Baldwin and Freeman and Carter had used (Scammon and Scott, 1927). Unaware of the readings, an observer planimetered the area of several circles which ranged from one to four centimeters in diameter 25 consecutive times. A second observer recorded each of the 25 readings. They determined that the Amsler planimeter failed to provide reliable readings. Therefore, they examined the Hammer area-by-weight method of determining ratios.

In the Hammer area-by-weight method, workers trace the desired area on special Hammer paper, meticulously cut out those areas and determine the desired ratio by weighting the two areas. To compute the Freeman-Carter ratio by the Hammer method, the workers would trace the carpal quadrilateral and the carpals. Next, they would cut out and weigh the carpal quadrilateral area, and then cut out and weigh the ossified carpal areas. By dividing the weight of the ossified carpal areas by the weight of the carpal quadrilateral area, the workers would determine the Freeman-Carter ratio.

However, Scammon and Scott realized that the Hammer method had three primary sources of error. (1) The ability of the workers to precisely cut out the elaborate outlines of the various structures varied. However, they found that careful operators minimized that source of error. (2) The paper's moisture content also varied, which caused the weight to vary depending on the ambient humidity. After much experimenting, Scammon and Scott decided that workers should use Eastman Kodaloid No. 3 celluloid in place of the Hammer paper. (3) The thickness of the Hammer paper also varied, which caused the weight to vary. With regard to uniform thickness, they determined that celluloid sheets varied 50 percent less than the Hammer paper. Because of the transparency of the celluloid, workers traced directly from the x-rays and since their sharp stilus easily cut through the celluloid, the Scammon-Scott method eliminated two steps from the Hammer method.

After Professor Bird T. Baldwin's death, Laura M. Bushby and Helen V. Garside wrote the final draft of Baldwin's ten year study on anatomic growth (Baldwin, Busby, and Garside, 1928). Iowa Child Welfare Research Station research associate S. Idell Pyle assisted Busby and Garside with the report's bibliography. Although Baldwin's study began in 1918, he collected the majority of his data between 1923 and 1926. Baldwin searched for basic anatomic growth principles, and if he uncovered a pattern, then he wanted to develop a general anatomic growth calendar.

After unsuccessfully experimenting with roller and polar planimeters, Baldwin et al. purchased Coradi's circumscribing disc planimeter to measure their area values. By remeasuring 56 carpals of eight ten-year-old boys, they determined that the reliability correlation coefficient of their planimeter measurements was .9996. To obtain their wrist rectangle measurement, they measured the width of the wrist from the most exterior points of the first and fourth metacarpals and the length of the wrist from the midpoint of the third metacarpal's proximal end to the midpoint of the radius' distal end, then they multiplied those two values. For the wrist rectangle measurements, they used a millimeter graduated sliding calipers.

Baldwin et al. conducted an experiment on the effect of x-ray hand position. By purposefully mispositioning their subjects' hands they found that the greatest difference in total carpal area was only 2.4 percent. Therefore, they concluded that children's ability to maintain the x-ray hand position had a negligible effect on the Anatomic Index value. Using several grades of paper and linen tracing cloth, they tried measuring carpal area and wrist rectangle by the Hammer weight method. However, they found that besides taking longer, the weight method decreased their accuracy. On the premise that the ratio between a carpal's area and its longest diameter remained constant, they measured each carpal's longest diameter. However, they found that the ratio of carpal area to longest diameter did not remain constant.

To compare the hand and wrist x-rays of boys and girls, Baldwin et al. recorded the age and order that they first observed each epiphyses, each carpal bone and each sesamoid bone, the age at the onset of ossification of the epiphyses, carpals and sesamoids, and the age that the distal ulnar and radial epiphysial fused. In every case, the girls' hands skeletally matured earlier than the boys' hands.

From the right and left hand x-rays of the 634 boys and 538 girls, Baldwin et al. recorded the number of carpal bones present, the area of each carpal and the total carpal area. From 200 older children's right and left hand x-rays, they determined: (1) the number of epiphyses present; (2) if those epiphyses had fused; (3) the number of sesamoid bones present; and (4) two wrist diameter measurements. To account for the relationship between carpal size and general skeletal size, they searched for an anatomic index. By dividing the total carpal area by the wrist rectangle, they computed an Anatomic Index. For both sexes at each age level, they determined a regression equation with carpal area the dependent variable and height and shoulder width the independent variables.

Using their Anatomic Index procedure, Baldwin et al. developed an Anatomic Index Table (Table 2.19). Except between the one to two

Table 2.19

Anatomic Index Values for the Boys and Girls Aged 1 to 17 Years (Baldwin, Busby, and Garside, 1928)

	Boys		Girls	
AGE (years)	N	Mean	N	Mean
1	19	.078	17	.082
2	16	.092	15	.108
3	22	.116	25	.144
4	26	.152	31	.180
5	31	.180	40	.248
6	40	.220	42	.314
7	45	.226	37	.373
8	48	.342	30	.401
9	31	.382	29	.474
10	30	.461	30	.529
11	40	.514	35	.598
12	42	.550	25	.630
13	35	.570	24	.689
14	37	.618	29	.685
15	38	.664	36	.707
16	24	.666	25	.706
17	18	.697	17	.712

year and 15 to 16 year old males, the 13 to 14, 15 to 16 and the 16 to 17 year old females, they found significant differences between every age level. Whereas the girls reached their maximum mean Anatomic Index by age 15, the boys still had not reached their maximum mean Anatomic Index by age 17.

In December of 1920, Physical Anthropology research associate Milo Hellman of New York's American Museum of Natural History x-rayed the right hand of 60 New York City Hebrew Orphan Asylum girls aged 123-147 months (Hellman, 1928). In the subsequent three years of his four year longitudinal study, he measured 56, 47 and 35 of those same girls. Hellman desired knowledge of the ages at which the hand epiphyses of girls ossified. Because he found its bones readily accessible and easily observable, he chose the hand as his skeletal development indicator. Each year he painstakingly recorded any morphological changes associated with the hand's ossification process. After several years of intense examination, he realized that he had to (1) delineate the various phases of epiphysial fusion, (2) arrange those phases in their manifested order, and (3) based on the girls' chronologic age, measure the onsets of those phases.

Using the median line, Hellman measured the millimeter length of each digit from the distal end of each first phalanx to the proximal end of its metacarpal. He also measured the millimeter length between the extreme ends of each phalanx, and between the end of the shaft and its diaphysis. As a result of his intense x-ray examinations, Hellman discovered that epiphysial ossification proceeded irregularly. Instead, he found that epiphysial ossification passed through five fundamental morphologic phases. Hellman labeled those phases Stage A through E and he stated that his five epiphysial ossification stages did not proceed sequentially (Table 2.20). Hellman also detailed the percent of the epiphyses that developed via different stages (Table 2.21).

Hellman concluded that epiphysial ossification started between 143 months and 155 months and terminated between 171 months and 180 months. Therefore, once started the entire hand epiphysial ossification process required less than 36 months.

Using the first three years data of a five year study by the New York City's Bureau of Educational Experiments, Anthropology

Table 2.20

The Five Stages of Epiphyseal Ossification

That Hellman Determined (Hellman, 1928)

Stages	Description
A	The epiphysis and diaphysis are distinctly separate.
B	The epiphysial cartilage narrows until the separation disappears.
C	A bud-like protrusion extends from the margins of the epiphysis and diaphysis. It can be one bud in the center and spread to the periphery, or, one or two buds at the periphery and spread towards the middle.
D	The actual fusion of the epiphysis with its diaphysis which shows on the X-ray as a dark line.
E	The epiphysis obtains the trabecular struction continuous with its shaft.

Table 2.21

The Percent of the Epiphyses that Developed

Via Different Stages (Hellman, 1928)

Transitions	Number	Percent
From A to B	419	30.6
From A to C	241	17.6
From A to D	26	1.0
From A to E	0	.0
From B to C	145	10.6
From B to D	127	9.3
From B to E	42	3.1
From C to D	195	14.2
From C to E	69	5.0
From D to E	107	7.8
totals	1371	100.1

Professor Ruth Otis Sawtell surveyed hand and wrist x-rays of children aged six months to seven years (Sawtell, 1929). She reported that whereas the children's carpal ossification was erratic during early childhood, their hand and wrist epiphysial ossification was orderly. She found that the ratio of the maximum transverse diameters of the epiphyses to its diaphysis provided consistent age and sex differences for children aged one to eight years old, and that the distal radial epiphysis to diaphysis ratio did not exceed 1.00 until after eight years of age. Therefore, by multiplying the radial epiphysis' maximum transverse diameter x 100, and then dividing by the distal radial diaphysis' maximum transverse diameter, Sawtell obtained her Radius Index measurement. Sawtell remarked that both she and Western Reserve University Professor T. Wingate Todd had selected the distal radial epiphyses and the metacarpal and phalangial epiphyses as the most promising aspects of the hand with which to solve the separate problems of differentiation uniformity and qualitative trait variability.

In her next report, with permission from Professor Todd, Professor Sawtell used her Radius Index to assess Western Reserve University's x-rays (Sawtell, 1929b). She compared her Bureau of Educational Experiment's children with Todd's Cleveland children (Table 2.22).

In 1930 Western Reserve University Professor T. Wingate Todd actively entered the research field of skeletal development. The first volume of the 1930 Child Development periodical contained three of Todd's articles on skeletal development.

Table 2.22

Sawtell's Comparison of the Bureau of Educational Experiments Sample and the Cleveland Sample (Sawtell, 1929)

Age(yrs)	Bureau of Educational Experiments Sample			
	Boys		Girls	
	N	Mean	N	Mean
1	10	17.4	7	24.9
2	15	43.4	11	55.7
3	16	59.2	14	70.3
4	21	66.2	24	79.4
5	22	76.7	24	87.5
6	20	83.0	21	91.0
7	11	91.6	15	95.5
8	10	95.0	10	99.8
Age(yrs)	Cleveland sample			
	Boys		Girls	
	N	Mean	N	Mean
5	9	79.3	9	89.1
6	40	83.8	31	89.3
7	40	89.3	32	92.9
8	44	94.0	30	99.3

Todd's first report contended that skeletal development specialists based their estimates of children's age of epiphysial union on insufficient valid evidence (Todd, 1930a). He said that basically their estimates came from Henle's 1891 report. In spite of impressive sample sizes, because x-ray researchers had never compared children's actual bones with x-rays of those bones, Todd questioned their observations. Additionally, Todd stated that because those x-ray workers failed to learn the skeletal maturation patterns of primates, they could not comprehend the skeletal maturation patterns of humans. Because Western Reserve University's

Hamann Museum housed more quantifiable human skeletons than any other institution, Todd said that no one could successfully refute his observations. The Hamann Museum possessed 200 children and young people skeletons, 150 anthropoid skeletons, and more than 850 mammalian skeletons.

Todd explained that although developing epiphyses form a nearly perfect cap, when this cap was large and complicated or deeply embedded in soft tissue, its x-ray easily confused the untrained worker. Therefore, he re-emphasized that except for those few x-ray workers that he had trained with actual skeletons, x-ray researchers could not relate the x-ray to the actual.

To evaluate the several stages of epiphysial union throughout the body, Todd examined the Hamann skeletons. After careful analysis he decided to label epiphysial unions as: (0) no union; (1) commencing union; (3) recent union; and (4) completed union. If the x-ray showed an indefinitely outlined whitish margin, then the actual bones showed slight bone tissue condensation between the diaphysis and epiphyses. When Todd compared the x-rays with the actual bones in his "commencing union" stage, he said that the actual bones showed the progressive narrowing of the diaphysoepiphysial gap. And he said that for a period of six months after epiphysial-diaphysial fusion, a fine red line marked the site of the union and this post-fusion red line eventuated to a white scar. Todd labeled the white scar stage as his "completed union" stage. By averaging the scores he assigned to every epiphyses in the skeleton, Todd estimated the skeleton's epiphysial maturation. Through careful dissections, Todd found that those epiphysial ossification stages occurred in other mammals as well as man.

Because he questioned the reliability of epiphysial union data of the standard anatomical textbooks, in 1923 Todd had asked his graduate student, P.H. Stevenson, to trace the origins of those data. Stevenson reported that except for Dwight's data in Piersol's Human Anatomy textbook, the standard anatomical textbooks based their data on Henle's 1871 report. Therefore, by examining the Hamann Museum's collection of over 200 children and adolescent skeletons, Todd compared the actual bones and the x-rays of those bones to find the facts about epiphysial union (Todd, 1930b). Because Todd discovered the epiphysial development proceeded in a complicated manner, and because previous skeletal maturation research lacked fundamental comparitive information, in spite of adverse reactions to Stevenson's 1924 report, Todd delayed his rebuttal until he clinically applied his observations. After analyzing the x-rays of over 1,000 living children and clinically applying his observations, Todd answered those opposed to Stevenson's 1924 report.

The x-ray sites where Todd said that workers could readily ascertain the successive stages of epiphysial development were the distal radial epiphysis and the distal humeral epiphysis. Todd described nine stages of epiphysial maturation (Table 2.23).

In his third article in the 1930 volume of Child Development, Todd discussed the principles of skeletal maturation and differentiation

Table 2.23

The nine epiphysial maturation stages that Todd described (Todd, 1930b)

Stages	Description
1	While a spherical epiphysial nodule progressively ossifies to reproduce the outline of the growing cartilaginous epiphysis, the diaphysis likewise ossifies.
2	Adjacent to the delimiting epiphysial and diaphysial surfaces, the bone texture changes. On the X-ray stage 2 shows a hazy ribbon between the surfaces.
3	The bony outlines develop a thin coating of more condenced osseous tissue. On the X-ray, Stage 3 shows a fine delimiting line of bone between the surfaces.
4	The bony outlines parallel each other. On the X-ray stage 4 shows a closing of the epiphysial-diaphysial gap.
5	The bony outlines parallel each other. On the X-ray stage 5 shows a definite parallelism in the clear billowy bony outlines.
6	The bony outlines narrow. Todd corresponded his stage 6 to Hellman's Stage B.
7	The bony outlines erratically extend across the epiphysial-diaphysial gap from the deeper surface towards the outer bone surfaces. On the X-ray, stage 7 shows as patchy irregularity of opposing surfaces with isolated double fine parallel lines. Todd corresponded his stage 7 to Hellman's Stage C.
8	On the outer surface of the bone the bony surfaces appear as a thin red line. On the X-ray, stage 8 shows a thicker single white line. Todd corresponded his stage 8 to Hellman's Stage D.
9	The osseous trabeculae extends from the shaft through the epiphysis and the bone loses all traces of epiphysial differentiation. Todd corresponded his stage 9 to Hellman's Stage E.

in the knees, elbows, and hands (Todd, 1930c). He noted that whereas epiphysial growth merely increased the bones' dimensions, when the shaft of the bones grew, the bones changed their proportion as well as their dimensions. As evidence that skeletal maturation disabilities involved only dimensional stunting and not differentiational stunting, Todd pointed to the number of short people versus the number of people with inhibited skeletal differentiation.

To select the X-ray standards for his skeletal maturation atlas, Todd catalogued his 1,000 hand-wrist X-rays by sex into six month age intervals. Then, within each group he ordered the X-rays from least to most differentiated. And finally, he selected the X-ray nearest the mean value as each group's standard. Thereafter, whenever they evaluated the skeletal age of their sample, Western Reserve University's child growth and development research team used Todd's atlas.

Todd acknowledged that Professor Sawtell had checked her Radial Index against his qualitative study, and that she had found both methods equally dependable. Therefore, when Todd discussed the principles of differentiation, he concluded that while the maturation process of the short bone shafts in the hand, the ulnar and radial epiphyses, and the carpals each uniquely described children's hand and wrist skeletal development, the distal radial epiphysis described hand-wrist skeletal development better than the rest.

In addition to his previous complaints about X-ray workers lacking training in reading X-rays properly and in understanding mammalian skeletal development, Todd argued that when those workers proposed a natural variability in skeletal development, they incorrectly analyzed the conditions under which those alleged variabilities occurred (Todd, 1931). When workers argued that the male and female skeletal maturation pattern differed, Todd countered that man and woman manifested a single genetic plan. While Todd found that female hands matured earlier than male hands, he also found that the female precosity did not proceed linearly (Table 2.24). Each group

Table 2.24

The female advantage over males in skeletal differentiation (Todd, 1931)

Ages (Years)	Skeletal Differentiation Advantage (months)		
	Knee	Hand	Elbow
5-0 to 5-5	6	0	0
5-6 to 5-11	6	0	0
6-0 to 6-5	6	6	0
6-6 to 6-11	12	6	0
7-0 to 7-5	15	6	0
7-6 to 7-11	12	6	0
8-0 to 8-5	9	6	15
8-6 to 8-11	12	4	9
9-0 to 9-5	6	0	6
9-6 to 9-11	18	0	6
10-0 to 10-5	18	12	24
10-6 to 10-11	24	15	24
11-0 to 11-5	18	18	18
11-6 to 11-11	12	12	12
12-0 to 12-5	12	12	9
12-6 to 12-11	6	4	24

had 20 to 25 subjects. Todd noted that at about nine years of age, the females' skeletal differentiation rate slowed until the sex difference disappeared for the hand and almost disappeared in the knee and elbow joints, but that after the tenth year, the females' skeletal differentiation rate accelerated again.

From 13 years of age upward, Todd evaluated children's skeletal development by epiphysial union (Table 2.25).

Studying the relation of race to skeletal maturation, Todd found that while white males and Negro males "punched the time clock together" and probably white and Negro girls also, American boys and girls skeletally differentiated a couple of months earlier than Italian boys and girls. After examining the relationship between children's bodily forms and rate of skeletal differentiation, Todd concluded that

Table 2.25

The female advantage over males in the age of epiphysial union (Todd, 1931)

AGE (years)	Epiphysial Union Advantage (months)
12-6 to 12-11	18
13-0 to 13-5	18
13-6 to 13-11	18
14-0 to 14-5	15
14-6 to 14-11	12
15-0 to 15-5	6
15-6 to 15-11	6
16-0 to 16-5	3
16-6 to 16-11	0

children with advanced height and weight displayed advanced skeletal development.

Eleven years after Prescott's monograph, Harvard University Professor Psyche Cattell re-examined Prescott's D/W ratio (Cattell, 1934). Cattell concluded that because both the numerator (D) and the denominator (W) increased with age, Prescott's ratio and any similarly designed index failed to provide satisfactory results.

To determine an objective ossification measurement, Cattell selected 54 annual X-ray plates of 9 feeble-minded girls from 12 to 18 years old. After inspecting those X-rays, Cattell decided on the following four measurement criteria: 1) the measurements should be objective, 2) the measurements should not be influenced by hand position changes during the X-raying procedure, 3) the measurements should be taken only on those bones that showed a prolonged growth period, and 4) the measurements should give equal weighting to the short bones, the long bones, and the epiphyses. Using those criteria, Cattell chose twelve measurements to begin her study (Table 2.26).

Table 2.26

The twelve hand-wrist measurements that Cattell used in her search for an objective assessment method (Cattell, 1934)

Number	Description of the measurement
1-5	The widest diameters of the Capitate, the Triangulate, the Lunate, the Havicular, and the Greater Multangular.
6-7	The lengths of the first and fifth metacarpal shafts.
8-9	The maximum transverse widths of the first and fifth metacarpal shaft and the proximal fifth metacarpal shaft.
10-11	The maximum transverse widths of the distal first metacarpal shaft and the proximal fifth metacarpal shaft.
12	The maximum transverse diameter of the distal radial epiphysis.

Because independent measurements by two regular Harvard Growth Study research assistants correlated at .95+, Cattell concluded that her twelve measurements not only compensated for random experimental errors, but also provided total annual increments sufficient to be reliable from birth through early adolescence.

Supported by a General Education Child Development Research Fellowship, University of Chicago's Charles D. Flory searched through anthropometric measurements and osseous indices to find a measure from which he could predict the age at which females reached puberty (Flory, 1935a). From the University of Chicago Laboratory Schools' Cumulative Records System, Flory selected 80 girls on whom they had taken yearly measures beginning with their tenth birthday. Flory chose to study females because the Child Research Center had more complete records on girls than on boys and because he could use the girls' first date of menstruation as a criterion of onset of puberty. Flory chose the following anthropometric measurements: 1) head width, 2) head length, 3) chest girth, 4) chest depth, 5) chest width, 6) lung capacity,

7) iliac width, and 8) trochanter width. The osseous indices that Flory chose were: 1) the Freeman-Carter Carpal Ossification ratio, 2) the age at Pisiforme appearance, 3) the age at the appearance of the sesamoid at the distal end of the first metacarpal, and 4) Flory's undefined skeletal age value. Flory determined that all of the osseous indices predicted the onset of female puberty better than any anthropometric measure, and that the girls' age at the time their first metacarpal appeared, coupled with their age when their Freeman-Carter ratio exceeded 1.00 provided the best predictor of the age at onset of female puberty.

From examining the right hand X-rays of 100 newborns, 300 one to four year old children, 6000 University of Chicago Laboratory School children, and 200 University of Chicago students, Charles D. Flory determined sex differences in the rate of skeletal growth and degree of skeletal maturation difference by the following four procedures (Flory, 1935b). 1) At annual intervals for each sex, Flory selected the X-ray typical of that group's skeletal maturation. When he compared the male and female standards, Flory found that the five year old female standard equalled the six year old male standard and the twelve year old female standard equalled the fourteen year old male standard. When he used epiphysial closure in the hand as his skeletal maturity standard, Flory found that fifty percent of the girls reached completed skeletal maturity by age seventeen, whereas fifty percent of the boys did not reach completed skeletal maturity until they were eighteen and one-half years of age. 2) Because the appearance order of the Naviculare, Lunatum, and Triquetrum carpals related inversely to their eventual size, their area growth curves had to cross. Therefore, by using the age at which those three area growth curves crossed as a measure of skeletal maturity, Flory found that the girls' Naviculare, Lunatum, and Triquetrum growth curves crossed at age eight and one-half years whereas the same curves in his male samples did not cross until the boys reached ten years of age. 3) By using total carpal area as a measure of skeletal maturity, Flory found that the girls neared their maximum total carpal area at fifteen years, whereas the boys did not near their maximum total carpal area until they were seventeen years of age. 4) By using the children's Freeman-Carter method as a measure of skeletal maturity, Flory found that the girls reached the 1.00 ratio just prior to age twelve years, whereas the boys did not reach the 1.00 ratio until almost fourteen years of age. Flory concluded that those four findings supported Pryor's contention that female hands matured earlier than male hands, first by days, then by months, and eventually by years.

In his 1936 Monograph, Charles D. Flory summarized the literature on children's hand-wrist development (Flory, 1936). In his opening statement, he presented: 1) the University of Chicago Laboratory School children's skeletal development norms, 2) his evaluation of the Freeman-Carter method, 3) his own skeletal development rating technique, and 4) a discussion of the predictability of several skeletal development measures.

As early as five years after the University of Chicago study that started in December of 1921, University of Chicago master degree candidate Paul M. Cook questioned the validity of the Freeman-Carter method. Flory cited the following three reasons why the Freeman-Carter method failed: 1) the slightest positional change during the X-raying procedure altered the resulting ratio, 2) because the first metacarpal epiphysis appears after one year of age for infant measures workers had to guess at that carpal quadrilateral landmark, and 3) when the children's skeletal maturity approached adult status, workers had problems locating the quadrilateral landmarks on the distal radial shaft and the distal ulnar shaft.

Because he determined that no ratio index could provide a completed skeletal development value, Flory developed an inspectional procedure for evaluating carpal development. As early as 1931, Flory supplemented the quantitative methods with his inspectional method. While Todd used the wrist as but one of several areas for which researchers could assess skeletal maturation, Flory explained that because he found X-raying several different joint areas and assessing those X-rays expensive and time consuming, he assessed only the wrist. Flory said that not only does the wrist X-ray easily, but complete skeletal maturation occurred in the wrist. Flory admitted that he had utilized many of Todd's suggestions.

For both sexes at annual increments between eight and seventeen, Flory selected 100 high quality X-rays. After arranging each group's X-rays from least to most skeletally mature, Flory chose the median X-ray as each group's standard. Then, Flory requested twenty graduate students and professors to independently order those standards from least skeletally mature to most skeletally mature. In every case, those untrained persons arranged Flory's standards as Flory had selected them. Along with the X-ray standard for each group and sex, Flory provided a written description of the skeletal maturation changes since the preceding standard.

One year after Flory's monograph, Professor T. Wingate Todd published his own *Atlas of Skeletal Maturation (Hand)* (Todd, 1937). Although Todd criticized Flory's work because Flory had studied only the hand and only hands of healthy youngsters, Todd did credit Flory with a useful research contribution to the study of skeletal maturation (Todd, 1937). Todd noted that during March of 1934, while he struggled with the verbal descriptions for his own standards, Dr. Flory had visited Todd's laboratory.

By 1937, Todd had already developed four separate sets of standards. He created his earlier 1927 and 1931 standards from chance selection of school children. But, because 1) Todd wanted age standard deviations of no more than three months for each standard, 2) his random sample selection had provided too few X-rays showing equal skeletal maturity throughout the skeleton, and 3) Todd had inadequate serial X-rays of children, Todd refused to publish either his 1927 or his 1931 standards. While his 1934 standards did have enough serial X-rays and sufficiently low age standard deviations, Todd still had reservations about the number of X-rays that showed unequal skeletal maturity throughout the skeleton. Even though his 1935 standards still had a few X-rays which displayed incongruent total body skeletal maturity, Todd decided to footnote his concern about those X-rays and publish his 1935 standards.

To create his 1935 standards, Todd and Dr. Carl E. Francis assessed over 4000 children's X-rays at least twelve times each, or almost one-half million separate X-ray readings. To select his age by sex standards, Todd divided the X-rays by sex into six month age groups. Within each of those groups, in addition to Dr. Francis and himself, Todd asked various assistants and Dr. H. V. Morley to use Todd's skeletal maturity indicators to order the X-rays from least to most skeletally mature. From the central mode of each group's X-rays, Todd chose the X-ray on which he had the least concern about unequal skeletal development. Todd said that because he had children of greater constitutional health in his 1927 and 1931 X-ray samples than in his 1935 X-ray sample, those early children's advanced skeletal maturation skewed his 1927 and 1931 standards towards accelerated skeletal maturity. Because of that skewness problem, as each group's X-ray standard, Todd selected each group's median X-ray as its standard. Only the time required by the trial and error method of practical experience delayed his presentation of those standards. Each of Todd's sex by age groups contained approximately 125 X-rays. Additionally, Todd argued that carpal and epiphysial appearance related to nutrition; therefore, they were not maturity indicators.

After her historical review of the measurement procedures for assessing hand and wrist skeletal maturation proved unrewarding, University of Iowa Child Welfare Research Station Professor Harriet J. Kelly decided to create her own skeletal maturation measurement procedure (Kelly, 1937). To select the most reliable and practical anatomic index, Kelly evaluated "four extent of ossification" measurements (Table 2.27) and five "size of the individual" estimates (Table 2.28) from the right hand and wrist X-rays of 703 Iowa City white males aged 69 to 219 months.

Table 4.27

Four "extent of ossification" measurements (Kelly, 1937)

Symbol	Description
CA	The disc planimetered sum of the carpal areas
CD	The sum of the carpal bones greatest diameters
Ca	The sum of Cattell's twelve measurements: the Capitate Hamate, Triangular, Lunate, Navicular, and Greater Multangular greatest diameters; the first and fifth metacarpal epiphyses broadest widths; the distal first metacarpal and proximal fifth metacarpal shafts broadest widths; and; the distal radial epiphysis widest diameter.
CM	The sum of the eight carpal bones greatest diameters, the distal radial and ulnar epiphyses widest diameters, and the first four metacarpal epiphyses broadest widths

Table 4.28

Five "size of the individual" estimates (Kelly, 1937)

Symbol	Description
WD	$\frac{1}{2}(A + B)$; the average wrist diameter
QA	$\frac{1}{2}(A + B) \times \frac{1}{2}(D + C)$; the quadrilateral area
DW	E; the diagonal of the wrist
QF	where $S = \frac{1}{2}(A+C+E)$ and $P = \frac{1}{2}(B+D+E)$, $QF = S(S-A)(S-C)(S-E) + P(P-B)(P-D)(P-E)$; determing the quadrilateral by formula
LM	The average length of the second, third, and fourth metacarpal bones

To calculate the estimates of the size of the individual, Kelly specified the following four points as her carpal quadrilateral landmarks: 1) the medial-most point on the first metacarpal shaft, 2) the medial-most point on the fifth metacarpal shaft, 3) the lateral-most point of the distal ulnar shaft, and 4) the medial-most point of the distal radial shaft. She defined Line A as the points between the first and fifth metacarpal shaft landmarks, Line B as the line between the distal radial and ulnar shaft's landmarks, Line C as the line between the fifth metacarpal shaft landmark and the distal ulnar shaft landmark, Line D as the line between the distal radial shaft landmark and first metacarpal shaft landmark, and Line E as the diagonal line between the first metacarpal shaft landmark and the distal ulnar shaft landmark. By combining those "four extent of ossification" measurements and five "size of the individual" estimates, Kelly created twenty indices.

When Kelly related those indices to chronological age, while the actual extent of the ossification increased, nine of their values decreased. Therefore, Kelly discarded those nine indices. Because she considered carpal area (CA) as a valid extent of ossification measure, and, because she had shown her wrist quadrilateral value as computed by her (QF) formula to be a valid estimate of the size of children's wrists, Kelly chose CA/CF as her evaluation criterion for the other indices. Kelly sought indices that required less measurement time; therefore, she eliminated four indices because they required the time-consuming task of planimetering carpal area. Because she believed that ratios of like quantities were more practical than ratios of unlike quantities, Kelly eliminated all of the mixed quantity ratios. Through those criteria, Kelly eliminated thirteen of her original twenty indices.

After applying the remaining seven indices to consecutive X-rays from twelve individuals, Kelly found that while the actual extent of ossification increased, only her Ca/QF and CM/WD ratio values did not decrease. Because 1) CM/WD correlated highly with her CA/QF criterion index and 2) to measure CM/WD utilized the highly reliable sliding caliper instead of the questionable reliability of a planimeter, Kelly selected the CM/WD ratio as her estimate of anatomic growth.

To test the CM/WD reliability, Kelly selected and assessed the following two experimental groups: 1) Group A contained 35 children aged 91 to 98 months and 2) Group B contained 35 children aged 177 to 182 months. After delaying six months, Kelly reassessed Groups A and B. By correlating her first and second CM/WD values, she found that their Pearson product moment coefficients were .99 and .98, respectively.

Using 710 X-rays of 151 girls and boys aged 3 to 60 months, Antioch, Ohio Fels Research Institute workers S. Idell Pyle and Camelle Menino compared the Flory monograph standards with the Todd Atlas standards (Pyle and Menino, 1939). To reveal their possible prejudices, the authors admitted that during the initial phase of Todd's work on his Atlas, Pyle had spent several months in Todd's laboratory, and after Todd had completed his Atlas, Menino had studied with Todd. Neither worker had met Dr. Flory.

On 3x5 notecards, Pyle and Menino independently rated the children's metacarpal shafts and bases; their proximal, middle, and distal phalanges; their carpals; their distal radial epiphysis; and their distal ulnar epiphysis. They always used Flory's monograph to assess the entire group first. To determine how nearly they agreed in their assessments, they randomly selected the skeletal age ratings for 75 children aged 36 months. Using Todd's Atlas and Flory's monograph for their assessments, they reported between raters Product-Moment correlation coefficients of .97 and .98, respectively. They attributed their higher Todd Atlas correlation to the fact that they practiced by rating with Flory's Atlas before they rated with Todd's Atlas, and to the fact that instead of Flory's annual increments, Todd's Atlas had standards every six months. On 50 children aged 36 months, they reported identical within-rater Product-Moment correlation coefficients of .87.

Because of the questionable value of appearance age of ossification centers to assess skeletal age, Pyle randomly selected and rated 50 girls aged 12 months (Table 2.29). With substitutes for six

Table 2.29

Pyle's Product-Moment correlation for fifty 12-month girls (Pyle and Menino, 1939)

Factors	Correlation
Todd's Skeletal Age and number of ossification centers	.82
Flory's Skeletal Age and number of ossification centers	.72
Todd's Skeletal Age and Flory's Skeletal Age	.91
With the number of ossification centers held constant, Todd's Skeletal Age and Flory's Skeletal Age	.81

girls they could not re-examine, Pyle assessed those same girls at age twenty-four months (Table 2.30). Tables 2.28 and 2.29 contain Pyle's correlation for those girls when they were 12 months old and when they were 24 months old. Because Pyle's 12-month to 24-month evaluations of Todd's skeletal age and number of ossification centers

Table 2.30

Pyle's Product-Moment correlation for fifty 24-month girls (Pyle and Menino, 1939)

Factors	Correlation
Todd's Skeletal Age and number of ossification centers	.67
Flory's Skeletal Age and number of ossification centers	.83
Todd's Skeletal Age and Flory's Skeletal Age	.90
With the number of ossification centers held constant, Todd's Skeletal Age and Flory's Skeletal Age	.83

correlations decreased from .82 to .67, they concluded that after toddlers' epiphyses differentiate, the number of ossification centers no longer influenced the Todd Atlas assessment. When they compared their ratings using the Flory monograph with ratings using the Todd Atlas, their Todd Atlas rating provided lower skeletal ages than their Flory monograph ratings. Pyle and Menino concluded that because Todd's Atlas had six month intervals for children aged six years or less, Todd's Atlas was superior to Flory's monograph.

Under the direction of Harold E. Jones, University of California Institute of Child Welfare research associate Nancy Bayley compared the central tendencies and standard deviations of her California sample with the norms of Todd's Atlas and Flory's monograph (Bayley, 1940). Except for the first three years of her study when she annually X-rayed the children, for eight years Bayley semiannually X-rayed the left hand and knee of children beginning at age eleven. Her California norms corresponded closely to Todd's and Flory's norms.

To test her measurement reliability, Bayley and another worker independently assessed the California children's X-rays twice. For those assessments of the California children's X-rays Bayley provided between and within raters correlations (Table 2.31). Because except for one .85 value she found that all of the correlations exceeded .90, Bayley hypothesized that their assessments reliably estimated skeletal age. To determine whether or not the hand and knee ratings differed, Bayley correlated her hand and knee ratings (Table 2.32). The high

Table 2.31

Between and within raters correlations (Bayley, 1940)

	Date	N	Hands Correlations	N	Knees Correlations
Within-Bayley	Spring 1935	90	.92	91	.92
	Fall 1935	87	.97	86	.94
	Spring 1936	86	.97	86	.96
	Fall 1936	82	.98	82	.94
	Spring 1937	78	.97	79	.94
Between Bayley	Fall 1937	79	.92	79	.91
and Rater "T"	Spring 1938	76	.94	76	.85
Within-Rater "T"	Spring 1938	76	.94	76	.90

Table 2.32

Hand and knee correlation table (Bayley, 1940)

	Boys		Girls	
Date	N	Correlations	N	Correlations
Spring 1935	90	.89	87	.91
Fall 1935	86	.89	84	.90
Spring 1936	86	.91	78	.93
Fall 1936	82	.87	because of the increased number of completed skeletal development ratings, Bayley stopped computing these correlation values.	
Spring 1937	78	.87		
Fall 1937	79	.90		
Spring 1938	76	.91		

correlations between the hand and knee reinforced Bayley's concept that skeletal development proceeded similarly throughout the skeleton. Nevertheless, Bayley concluded that hand assessments estimated skeletal age better than knee assessments. However, when studying the relationship between skeletal development and other data, she advocated that a maturity score from the combined hand and knee assessments estimated skeletal age better than any assessment of either area alone.

Although in most instances her cross-sectional assessments remained unchanged, to check the accuracy of isolated cross-sectional assessments, Bayley suggested that workers serially order the children's X-rays and use those longitudinal assessments to compare with their cross-sectional assessments. Bayley contended that when researchers examine children's X-rays longitudinally, those researchers more readily notice peculiarities in individual bone contours and chance variations in bone positions, and thereby they make fewer assessment errors.

Influenced by the late T. Wingate Todd, the Brush Foundation averaged the hand and wrist, elbow, shoulder, hip, knee, and foot and ankle assessments. Advised by Dr. W. W. Greulich, Western Reserve University Brush Foundation workers Clara C. Beuhl and S. Idell Pyle searched for the minimum number of those six areas that would provide a reliable estimate of skeletal status (Beuhl and Pyle, 1942). By comparing the ages at first appearance of the distal ulnar epiphysis, the thumb's first sesamoid, and the iliac's crest, they tested the value of onset of ossification as a skeletal maturity determinator. To reduce the required number of X-rays, they substituted the onset of epiphysial fusion of the index finger's distal phalanx for the onset of ossification in the iliac crest. Because onset of ossification in the iliac crest of females occurred at the time of menarch, they studied whether or not the onset of ossification in the iliac crest of males also indicated the beginning of their reproductive ability. Todd believed that the onset of epiphysial calcification was particularly sensitive to environmental factors, and therefore, except for extremely healthy children, he advised against using onset of epiphysial calcification as a maturity indicator. However, from their more recent evidence, Beuhl and Pyle concluded that epiphyses were no more sensitive to their environment at onset of calcification than at any other period in their development.

Brush Foundation workers X-rayed their skeletal development sample every three or six months until five years of age and annually thereafter. Because 1) the distal ulnar epiphysis, the humeral olecranon process, the humeral medial epicondyle, and the calcaneus epiphysis begin to calcify at about five years of age, 2) they wanted a skeletal development measure at the five year old age level, and 3) they wanted to decrease the number of required X-rays, Beuhl and Pyle selected the distal ulnar epiphysis as their skeletal development criterion for the 5½ to 7½ year olds. They labeled this age range as the Ulnar Age. Because the thumb's first sesamoid bone begins to calcify before the onset of adolescence, Beuhl and Pyle selected the thumb's first sesamoid to represent the 10 to 12 year old age level. They labeled this age range as the Sesamoid Age. Then, they selected the iliac crest to represent the 13 to 14 year old age level. They labeled this age range as the Iliac Age.

Using 10 annual X-rays of 30 boys and 30 girls starting at age five, Beuhl and Pyle interpolated the ages at ossification onset for the distal ulnar epiphysis, the thumb's first sesamoid, and the iliac crest. By comparing Ulnar Age, Sesamoid Age, and Iliac Age with total skeletal age, except for the girls' Sesamoid Age, they found no significant differences between the mean total skeletal age measures and the Ulnar Age, Sesamoid Age, and Iliac Age measures. And, they found that whereas the mean total skeletal age correlated .24 with menarchial age, the distal ulnar epiphysis correlated .62 with menarchial age. Therefore, they concluded that while a correlation of .62 insufficiently predicted female puberty, considering the Ulnar Age preceded menarch by seven years, the age at ossification for the distal ulnar epiphysis significantly exceeded the predictive ability of mean total skeletal age. Also, they found that whereas the mean total skeletal age correlated .58 with menarchial age, Sesamoid Age correlated .71 with menarchial age, and whereas the mean total skeletal age correlated .50 with menarchial age, Iliac Age correlated .83 with menarchial age. In their conclusions, they stated that Ulnar Age, Sesamoid Age, and Iliac Age were as reliable as mean total skeletal age and, Ulnar Age, Sesamoid Age, and Iliac Age correlated with menarchial age better than mean total skeletal age.

Comparing the age at onset of epiphysial fusion in the index finger's distal phalanx with the onset of ossification in the iliac crest, Beuhl and Pyle found that whereas for boys the onset of ossification in the iliac crest started before epiphysial fusion in the index finger's distal phalanx, for girls the epiphysial fusion of the index finger's distal phalanx started before the onset of ossification in the iliac crest. Nevertheless, because they wanted to decrease the number of required X-rays, Beuhl and Pyle substituted the age of onset of epiphysial fusion in the index finger's distal phalanx for the onset of ossification in the iliac crest.

At the *Symposium on the Concept of Maturity from the Anatomical, Physiological, and Psychological Points of View* sponsored by the Society for Research in Child Development and held in New York City on December 28, 1949, Stanford University Professor William W. Greulich presented his rationale for using hand and wrist X-rays to assess children's developmental status (Greulich, 1950). Greulich said that because the genetic heritage of American children varied considerably, he found that chronological age, height, and weight were not reliable indicators of growth and development.

To illustrate the relationship between skeletal development and sexual maturation, Greulich provided both hand-wrist X-rays and photographs of a girl with precocious puberty aged 23 to 72 months, and of a twenty-two year old hypogonadal male. Greulich determined that at 72 months of age the girl with precocious puberty had a skeletal age and sexual maturation equal to a normal girl aged 172 months. Contrarily, the twenty-two year old hypogonadal male had the skeletal age and sexual maturation equal to that of a normal sixteen year old male.

By arranging the mean annual height gains of seven menarchial age groups of normal girls so that their menarchial age point touched the same vertical line, Greulich demonstrated that the female's maximum annual height increment occurred during the year preceding menarch. Greulich said that because healthy, adequately nourished children's skeletons develop as a unit, the skeletal maturation of their hands and wrists are valid indicators of total skeletal maturation. Because of that relationship between the maturational changes in the skeletal and reproductive systems, Greulich used hand-wrist X-rays to predict menarch. He found that menarch occurred between the onset and completion of phalangial union, usually immediately after distal phalangial epiphysial fusion.

Greulich also used X-rays to predict early versus late maturing children. For normal boys, he found that changes in their primary and secondary sex characteristics accompanied or followed their maximum annual height gains. Greulich provided photographs of 10 boys aged 168 to 170 months. While he observed that the development of their primary and secondary sex characteristics ranged from early prepubertal to slightly sub-adult, Greulich determined that from 148 months to 200 months their skeletal ages ranged in direct relation to their sexual maturation. Greulich contended that the hand-wrist X-rays of those 10 boys demonstrated the close relationship between the developmental status of the reproductive and skeletal systems. Therefore, the degrees of development of the reproductive and skeletal systems reliably indicated the boys' general bodily maturity. However, because the skeleton reflected the body's general level of maturation from birth to completed skeletal development, Greulich said that the hand-wrist X-ray provided a better biological maturation estimate than primary and secondary sex characteristics.

Greulich provided the following five reasons why he believed that hand-wrist X-rays determined developmental status better than sexual maturation: 1) hand-wrist X-rays readily showed progressive maturity changes, 2) hand-wrist skeletal maturation occurred in an orderly sequence, 3) hand-wrist skeletal maturation covered the entire period from birth to early childhood, 4) without regard to genetic differences, hand-wrist X-rays permitted direct comparisons between children, and 5) hand-wrist skeletal maturation was intimately related to sexual development.

As a sequal to Todd's 1937 *Atlas of Skeletal Maturation (Hand)*, Stanford University Anatomy Professor William W. Greulich (former Brush Foundation Director) and Brush Foundation research associate S. Idell Pyle had the Stanford University Press publish their *Radiographic Atlas of Skeletal Development of the Hand and Wrist*. In the book's preface they explained how in 1931 T. Wingate Todd had started the Brush Foundation's twelve year investigation of human growth and development (Greulich and Pyle, 1950). They noted that pediatricians had recommended the early Brush Foundation children, and that those children were of white North European ancestry with above average economic and educational status. To complete his 1937 Atlas, they said that Todd included children from various Cleveland public schools and social agencies. The Greulich and Pyle Atlas included only children in the Brush Foundation sample. Other than the different sample, the Greulich-Pyle Atlas differed from Todd's Atlas only in that they provided annual standards from age five to age thirteen in girls and from age five to age fifteen in boys. However, to accommodate the rapid change in skeletal age during the adolescent years, they did include a standard for the thirteen and one-half year old female and a standard for the fifteen and one-half year old male.

Because they discovered that children suffering repeated illnesses displayed incongruencies in their skeletal maturation, Pyle developed a graphic method for describing skeletal status. By plotting the children's serial ages of individual hand-wrist bones on graphs, Pyle's procedure disclosed the fluctuations between the least and most skeletally mature bones. Therefore, until the X-ray readers devised their own method for assessing the X-rays, Greulich and Pyle suggested a six-step procedure (Table 2.33).

Table 2.33

Suggested procedure for X-ray assessment according to Greulich and Pyle (Greulich and Pyle, 1950)

Step	Description
1	Compare the X-ray with the Atlas' standard of the same sex and nearest chronological age.
2	Compare the X-ray with the Atlas' next oldest and youngest standard.
3	For a more detailed comparison, select the standard that generally resembles the X-ray.
4a	If the X-ray is of infant to early childhood children, then the presence of carpal and/or epiphysial centers provide the guidelines.
4b	If the X-ray is of later childhood to pre-pubertal children, then the shape of the carpals and our list of maturity indicators provide the guidelines.
5	To prevent overlooking features, develop a regular order for assessing the individual bones and epiphyses; such as; distal radius, distal ulna, carpals, metacarpals, and phalanges. View the carpals in a regular order; such as; Capitate, Hamate, Triquetral, Lunate, Navicular, Greater Multangular, Lesser Multangular, and Pisiform.
6	For each bone assign the age of the standard selected that resembled it. If the bone's development falls between two standards, estimate the age at which the bone has progressed from the lower standard to the higher standard.

To enable the physician to determine the extent to which previously ill children had corrected their imbalanced skeletal development, they suggested that the X-ray readers plot the least and most mature hand-wrist osseous center. The smaller the difference between the skeletal maturation of those osseous centers, the closer the children were to recovery from their illness. During their discussion of what constituted a significant deviation from normal in skeletal age relative to chronologic age, they provided a table of the variability in skeletal age of the boys and girls in the Brush Foundation study (Table 2.34). If Greulich and Pyle had used their

Table 2.34

Variability in the skeletal age of the boys and girls in the Brush Foundation Study (Greulich and Pyle, 1950)

Number	Chronological Age (months)	Skeletal Age (months)
121	3	3.01
129	6	6.09
137	9	9.56
130	12	12.74
106	18	19.36
105	24	24.97
107	30	32.40
127	36	38.21
138	42	43.89
170	48	49.04
176	54	56.00
191	60	62.43
186	72	75.46
182	84	88.20
168	96	101.35
160	108	113.90
177	120	125.68
154	132	137.32
165	144	148.82
175	156	158.39
163	168	170.02
124	180	182.72
99	192	195.35
68	204	206.21

Atlas to assess the skeletal ages of the Brush Foundation sample, then the mean skeletal ages should have equalled the chronological ages. However, because they used Todd's Atlas, the Greulich-Pyle skeletal ages exceeded those that Todd reported. Therefore, because Greulich and Pyle used the Brush Foundation children exclusively for their 1950 Atlas, their standards exceeded those in Todd's 1937 Atlas.

In the back of their Atlas, Greulich and Pyle provided detailed descriptions and schematics of the maturity indicators for their twenty-one hand-wrist bones (Table 2.35). For each hand-wrist

Table 2.35

Number of maturity indicator stages for the twenty-one hand-wrist bones (Greulich and Pyle, 1950)

Number	Hand-Wrist Bones	Number Maturity Indicator Stages
1	Distal Radial Epiphysis	13
2	Distal Ulnar Epiphysis	10
3	Capitate	8
4	Hamate	10
5	Triquetral	8
6	Lunate	7
7	Navicular	8
8	Greater Multangular	10
9	Lesser Multangular	8
10	1st Metacarpal Epiphysis	9
11	2nd Metacarpal Epiphysis	9
12	2nd and 3rd Metacarpal Epiphyses	8
13	4th and 5th Metacarpal Epiphyses	8
14	Thumb's Proximal Phalanx Epiphysis	8
15	2nd and 3rd Fingers Proximal Phalange's Epiphyses	9
16	4th and 5th Fingers Proximal Phalange's Epiphyses	9
17	2nd and 3rd Fingers Middle Phalange's Epiphyses	8
18	4th and 5th Fingers Middle Phalange's Epiphyses	8
19	Thunb's Distal Phalanx Epiphysis	8
20	2nd and 3rd Fingers Distal Phalange's Epiphyses	8
21	4th and 5th Fingers Distal Phalange's Epiphyses	8

bone, the number of maturity indicator stages varied. They warned that because 1) their maturity indicators were not equivalent or of the same relative value and 2) hand-wrist X-ray position could distort the maturity indicators' appearance, assessors of X-rays should only use the maturity indicator sketches as auxiliary guidelines.

University of California Institute of Child Welfare Professors Nancy Bayley and Samuel R. Pinneau found that their X-ray readings from the 1937 Todd Atlas and the 1950 Greulich-Pyle Atlas differed (Bayley and Pinneau, 1952). Therefore, because they expected the 1950 Greulich-Pyle Atlas to replace the 1937 Todd Atlas, they revised Bayley's 1946 Tables for Predicting Adult Height from Skeletal Age that she had based on the 1950 Greulich-Pyle Atlas. When they completed their new tables and compared them with their 1946 tables, they concluded that the two sets of tables compared favorably.

University of Oxford Professor Roy M. Acheson provided the following four reasons why he objected to the inspectional techniques offered by the Todd 1937 Atlas and the Greulich-Pyle 1950 Atlas: 1) whereas considerable evidence has shown a wide variability in ossification patterns, the Todd and Greulich-Pyle Atlases presupposed a fixed pattern of appearance time and order and subsequent development of the hand-wrist carpals and epiphyses; 2) because when the bones fit between two standards, the raters of X-rays had to interpolate the bone age, the six month intervals between the Todd and Greulich-Pyle standards were too large, and, therefore, the raters of X-rays unavoidably made assessment errors; 3) because the Todd and Greulich-Pyle Atlases had separated the standards for each sex, raters of X-rays had to further qualify terms like "skeletal age 30 months" as to whether the subject was male or female; and 4) because the skeletal maturation process does not follow the same schedule as chronological age, just as height and weight have their own dimensional units, skeletal maturation should have its own dimensional unit (Acheson, 1954).

To devise his method of assessing skeletal maturation, Acheson set up the following three criteria: 1) regardless of the children's ossification pattern, because to each X-ray assessment all round bones and epiphyses make their own contribution, the raters of X-rays can evaluate all X-rays equally; 2) raters of X-rays should record even the small increases in skeletal maturation; and 3) because raters of X-rays will record skeletal maturation in Oxford Maturity Units, they will use the same standards for both sexes.

Acheson decided that as each bone's maturity indicator stages appeared, the raters would assign that stage as the bone's Oxford Maturity Unit. Therefore, the sum of all of the hand-wrist bones' Oxford Maturity Units represented an exact measure of skeletal maturity. As his basis for determining the hand-wrist maturity indicator stages, Acheson selected the descriptions and illustrations that Greulich and Pyle reported.

In 1959 Greulich and Pyle revised their 1950 Atlas (Greulich and Pyle, 1959). They improved the reproduction of their standards, and where they determined that the intervals between standards had been too long, they added new standards. Additionally, they revised their descriptions of the maturity indicators and had their maturity indicator schematics redrawn. In a new appendix section, Greulich and Pyle added the Bayley and Pinneau 1952 Tables for Predicting Adult Height from Skeletal Age.

For the purpose of collecting information on the health of the U.S. population, in 1956 the Congress authorized the U.S. Public Health Service to conduct a National Health Survey on a continuing basis. Because they determined that children's growth and development were legitimate health data, in 1963 the staff of the National Center for Health Statistics (NCHS) started assessing children's skeletal maturation. In addition to being an important basis for classifying children's health data, they believed that a representative current sample of U.S. children provided reliable normative data on children's skeletal development.

To determine the available hand-wrist X-ray assessment procedures and the uses thereof, the NCHS staff requested advice from clinicians and directors of long-term skeletal maturation studies. Among the directors that the NCHS staff contacted were Harvard Growth Study Director Harold C. Stuart and former Brush Foundation Director William W. Greulich. Stuart and Greulich recommended that the NCHS staff ask Case Western Reserve University Developmental Anatomy research associate S. Idell Pyle to prepare the National Health Survey's hand-wrist X-ray standards. By 1962 specialists in child growth and development research had determined that they needed only one series of hand-wrist X-rays for both sexes, provided they assigned male and female chronological age equivalents to each standard.

NCHS's Chief Medical Advisor Alice M. Waterhouse and co-worker Tavia Gordon found Professor Pyle already preparing standard X-ray references for children's shoulder, elbow, and hand-wrist joints. Using subjects when she determined had moderate skeletal maturation rates, Pyle had assigned male-female age equivalents to her standards. As Professor Todd had done, Pyle decided to use semi-annual standards between the ages of three and thirteen years.

Between July 1963 and the end of 1965, by using four specially constructed and equipped tractor trailers, the NCHS staff examined 7,100 children ranging in age from 6 to 12 years (Pyle, Waterhouse, and Greulich, 1971). They parked the trailers in forty locations across the United States. Then, from 1966 to 1970, the NCHS staff examined children aged 12 to 18 years, thirty percent of whom they had assessed in their 1963-1965 tour. To design a hand-wrist X-ray reference Atlas, the NCHS staff of Paul T. Bruyere, Arthur J. McDowell, Alice M. Waterhouse, Tavia Gordon, Elizabeth S. Johnson, Peter V. Hamill, and S. Idell Pyle made the five assumptions about the ossification process (Table 2.36).

Table 2.36

Assumptions of the human ossification process that provided the bases for the NCHS Atlas of hand-wrist maturity (Pyle, Waterhouse and Greulich, 1971)

Number	Description
1.	In the cortexes of growing bones, the transitions, ossifying features and articular facts are the same for both sexes and across all races.
2.	During both childhood and adolescence, the chronological intervals of the phases through which growing bones develop vary between very brief and prolonged.
3.	In males and females on the same skeletal development schedule, the ossification process occurs more slowly in males.
4.	For direct skeletal age readings, a single standard X-ray series must show male and female bone ages.
5.	When using the single standard X-ray bone age series for only one sex, although the skeletal age reading for the opposite sex is an indirect reading, nevertheless, we can translate that indirect reading into its proper sex-related skeletal age.

The clinicians and directors that the NCHS staff consulted proclaimed that they had tested the NCHS's first three assumptions and had found those assumptions to be correct. With regard to the NCHS's fourth assumption, the NCHS staff found that Professor Pyle and her associates had tested that assumption twice, and based upon the accuracy of that assumption, they had published their 1962 *Radiographic Atlas of Skeletal Development of the Foot and Ankle* and their 1969 *Radiographic Standard of Reference for the Growing Knee*. The clinicians and directors said that they did not know if anyone had tested the NCHS's fifth assumption.

In addition to the Flory Monograph, the Todd Atlas, and the Greulich-Pyle Atlas, the NCHS staff examined B. Speijer's 1950 Atlas that he based on hand-wrist X-rays of Dutch children, D. W. MacKay's 1952 Atlas that he based on the hand-wrist X-rays of East African children, and W. W. Sutow's and K. Ohwada's 1953 Atlas that they based on the hand-wrist X-rays of Japanese children. When constructing those Atlases, Greulich and Pyle, Speijers, MacKay, and Sutow and Ohwada admitted that they referred to Flory's Monograph and Todd's Atlas. Consequently, although the six Atlases provided separate X-ray standards for males and females, nevertheless, of those workers, all designed their X-ray standards to enable the assessors to determine the skeletal age for individual hand-wrist bones without regard to sex differences.

After agreeing that sufficient empirical and experimental evidence supported their first four assumptions, in 1965 the NCHS staff gave their hand-wrist X-ray assessors the following three directives: 1) you will not know the name, age, or sex of the person whose X-ray you are rating; 2) using the 1959 Greulich-Pyle male standards only, you will assign a bone age to each hand-wrist bone; and 3) when a bone does not match a standard perfectly, you will interpolate its bone age mathematically.

By March 1971, eleven Case Western Reserve University medical school students that S. Idell Pyle had trained in hand-wrist X-ray rating had assessed the skeletal ages of 377,800 hand-wrist bones from 13,496 X-rays. From that work, the NCHS provided the equivalent hand-wrist skeletal ages for males and females (Table 2.37).

Table 2.37

The equivalent hand-wrist skeletal ages (months) for males and females (Pyle, Waterhouse, and Greulich, 1971)

Female	Male	Female	Male	Female	Male	Female	Male
50	36	91	78	143	120	193	162
55	42	97	84	153	126	198	168
61	48	104	90	163	132	203	174
68	54	111	96	170	138	207	180
74	60	117	102	176	144	211	186
80	66	123	108	182	150	216	192
85	72	133	114	187	156	228	206

While having twelve standards for females with skeletal ages less than 5 years, the NCHS Standard of Reference for the National Health Examination Survey provided twenty-six hand-wrist standards for children ranging in age from 3 to 228 months. On the page beside each plate, the NCHS Standard of Reference listed the sex age equivalents for each bone, and it discussed the bone ossification that had occurred since the preceding plate.

Summary

In the history of the assessment of hand and wrist skeletal development, the author discussed the research that American specialists in skeletal development had conducted beginning with Professor Pryor's work in 1905. The discussion traced the research from early works of Pryor and Rotch through the University related children research centers phase that started around 1920 up to the domination of the field by T. Wingate Todd beginning in 1930. The early hand-wrist skeletal development research encompassed many different inspectional and measurement methods. However, when Professor T. Wingate Todd entered the skeletal maturation research field, the concept of assessing skeletal maturation by non-inspectional methods diminished. Although Todd rushed his 1937 Atlas into print, his concepts have continued. Because of Todd's unexpected death in 1938, William W. Greulich and S. Idell Pyle have carried on Todd's work.

From all of this work, some of the skeletal maturation principles pertinent to this investigation are: 1) the skeletal maturation phenomena represent children's true biological age better than any other measurable phenomenon of growth and development; 2) because of the presence of two long bone epiphyses, eight carpal bones, five metacarpal epiphyses, and fourteen phalangial epiphyses, skeletal maturation estimates based on the hand and wrist reliably represent the total skeleton; and 3) the 1959 Greulich-Pyle hand-wrist Atlas provides a reliable criterion upon which assessors of hand-wrist X-rays can base their estimates of children's skeletal age.

HISTORY OF THE ASSESSMENT OF ADOLESCENT MALES' SEXUAL CHARACTERISTIC DEVELOPMENT

In 1900 New York City Physician C. Ward Crampton measured the physical and mental development of adolescent males in an attempt to compare the onset of puberty in males with the onset of menarch in females (Crampton, 1904). From his vast measurement experience, Crampton believed that the best available indicator of the onset of male puberty was the boys' pubic hair development. After considerable cataloging, Crampton decided that pubic hair development had three distinct stages (Table 2.38). Because of insufficient data, Crampton initially divided the pubescent stages into two sub-groups.

Table 2.38

Pubic hair stages (Crampton, 1904)

Classification	Description
1. Pre-pubescent	No pubic hair
2. Pubescent	a. straight, unpigmented pubic hair
	b. straight, pigmented pubic hair
3. Post-pubescent	Kinked and curled, pigmented pubic hair

From data he collected on 1,200 male students aged 11 to 18 years from New York City High School of Commerce, Crampton determined the percentage of boys that he classified in his three stages of pubic hair development. Including only the boys aged 13½, 14½, and 15½ years, he provided the percentage of boys at each age that he rated in each stage (Table 2.39).

Crampton noted the orderly pattern of the boys' pubic hair development. By multiplying the percentage of boys at a specific age by the number he had assigned to the pubescent stage, Crampton created

Table 2.39

Percentage of boys at each age that Crampton rated in each stage (Crampton, 1904)

Pubescent Stage	Age		
	13½	14½	15½
1	51%	23%	12%
2	29%	34%	22%
3	18%	42%	65%

his "Puberty Index". For example, by using the values shown in Table 2.38 for the 13½ year old boys, by multiplying (51 x 1, 29 x 2, and 18 x 3), we can determine that the "Puberty Index" for the 13½ year old boys is 51+58+54 = 163. Using his Puberty Index, Crampton compared boys whose parents were born in the United States with boys whose parents were born in Germany (Table 2.40). Crampton concluded that boys whose parents were born in Germany entered pubescence later than boys whose parents were born in the United States.

Table 2.40

The "Puberty Index" of boys with United States versus German heritage (Crampton, 1904)

Heritage	Ages		
	13½	14½	15½
U.S. born parents	165	219	248
Germany born parents	150	195	219

To study variability in the boys' height, weight, and right hand grip strength, Crampton divided the boys by his Pubic Hair Index within each chronologic age level (Tables 2.41, 2.42, and 2.43). Crampton noted that at every age level the taller, heavier, and stronger boys had the more advanced pubic hair development.

Table 2.41

The mean values for height at each age for each pubic hair stage (Crampton, 1904)

Age	Pubescent Stage	N	Mean
13½	1	109	145.52
	2	54	148.12
	3	40	155.25
14½	1	77	146.26
	2	107	148.17
	3	130	159.38
15½	1	27	148.79
	2	44	152.61
	3	131	161.31

Table 2.42

The mean values for weight at each age for each pubic hair stage (Crampton, 1904)

Age	Pubescent Stage	N	Mean
13½	1	116	39.92
	2	53	39.32
	3	49	41.48
14½	1	76	35.53
	2	108	39.02
	3	132	47.50
15½	1	28	38.38
	2	45	39.38
	3	132	49.54

Table 2.43

The mean values for right hand grip strength at each age for each pubic hair stage (Crampton, 1904)

Age	Pubescent Stage	N	Mean (Kilos)
13½	1	93	28.15
	2	53	32.24
	3	34	35.88
14½	1	67	28.28
	2	94	31.28
	3	115	41.44
15½	1	23	29.78
	2	40	31.75
	3	125	43.24

After collecting data for four additional years, C. Ward Crampton reported his findings on 3,835 New York City High School boys aged 11 to 18 years old (Crampton, 1908a). In half year intervals, he provided the percentage of boys that he rated in each of his three pubic stages (Table 2.44). Because he did not have any subjects less

Table 2.44

Percentage of boys in the three pubic hair stages at six month intervals (Crampton, 1908a)

Pubertal Stage	Age (months)											
	147	153	159	165	171	177	183	189	195	201	207	213
1	(81)	69	55	41	26	16	9	5	2	1	0	0
2	(16)	25	26	28	28	24	20	10	4	4	2	0
3	(3)	6	18	31	46	60	70	85	93	95	98	100

than twelve years old, Crampton used parentheses to denote that he had estimated the percent in the 147 month age group. Because the largest increments of maturation occurred between ages of 165 and 171 months and 183 and 189 months, and the smallest increments occurred between the ages of 177 and 183 months, Crampton concluded that as measured by his Pubic Hair Index, adolescent males' physiological maturation had a bimodal distribution centered on the age of 174 months. By observing the rate at which the 69 percent of pre-pubescent boys at age 153 months advanced to their pubescent pubic hair stage, Crampton examined the variability in adolescent males' rate of physiological maturation. He noted that while 14 percent of those boys became pubescent within six months, he still rated 17 percent of them pre-pubescent twenty-four months later. In addition to his use of presence and type of pubic hair as his criteria for pubescence, Crampton suggested that the amount of pubic hair, the appearance and amount of axillary hair, and the appearance and amount of hair on the upper lip could also serve as useful criteria for assessing pubescence.

From an unsolicited source, Crampton received research evidence of well-formed spermatozoa in the secretion of the boys he rated as pubescent. After inspecting the spermatozoa-pubic hair relationship, Crampton determined that those pubescent boys with urine containing well-formed spermatozoa were within a few months of his post-pubescent rating. Crampton recommended that someone should conduct a thorough investigation on the relationship of his pubic hair stages to the presence of well-formed spermatozoa.

In a second 1908 article, Assistant Director of New York Public Schools, C. Ward Crampton explained that he created his second stage of pubic hair development because raters of pubic hair development had some trouble differentiating between those boys with only some pubic hair and those boys with an abundance of pubic hair (Crampton, 1908). Therefore, whenever raters were uncertain about the boys' pubic hair stage Crampton suggested that they rate those boys as pubescent. As had Thomas Morgan Rotch with his Anatomic Index, so Crampton used his Pubic Hair Index to advise educators and employers of the importance that children's biological age rather than their chronological age had on their readiness to meet educational or vocational responsibilities.

Twenty years later at the University of Iowa Child Welfare Laboratory, University of Iowa Professor Bird T. Baldwin studied the relationship between the development of sex characteristics in males and their production of spermatozoa (Baldwin, 1928). Between the dates of May 23, 1923, and April 5, 1924, Baldwin collected 1,135 6:00 a.m. urine samples from 123 state reform boys ranging in age from 9 to 17 years. The University laboratory tested those samples for the presence of spermatozoa. If they found a positive slide, they stained the spermatozoa with carbol fuchsin, and photographed the slide. Baldwin provided the frequency of spermatozoa that they found at each annual age level (Table 2.45). They found spermatozoa

Table 2.45

The number of positive spermatozoa samples found at each age level (Baldwin, 1928)

Descriptive Data	Ages (Years)								
	9	10	11	12	13	14	15	16	17
Number of Urine Samples	10	30	64	231	206	188	305	75	27
Number of Positive Specimens	2	6	10	29	21	20	25	8	2

in 29 of the 123 state reform school boys. Baldwin examined 24 of the 29 boys with positive spermatozoa specimens for axillary and pubic hair development. He found 18 boys had axillary hair pigmentation and pubescence and 23 boys had pubic hair pigmentation and pubescence.

To establish more definite standards for the physiological development of primary and secondary sexual characteristics in adolescent males, and to establish the relationship between their testicular development and the development of the secondary sexual characteristics, University of Pennsylvania Graduate School of Medicine Physician Paul E. Kubitschek assessed the primary and secondary sexual characteristics of 730 boys aged 9 to 18 years (Kubitschek, 1932). Although he excluded all Hebrew boys from the study, Kubitschek included 125 Negro boys in his cross-sectional data collection. Kubitschek developed a sexual developmental checklist that he used to assess the boys' sexual development (Table 2.46). Kubitschek

Table 2.46

The sexual development checklist that Kubitschek developed (Kubitschek, 1932)

Sexual Chacteristic	Evaluation						
Penis	Length_________cm.			Thickness____________cm.			
Scrotum							
Testicles	Peanut		large medium small	Almond	large medium small	Walnut	large medium small
-							
Pubic hair	0	1	2	3	4		
Axillary hair	0	1	2	3	4		
Chest hair	0	1	2	3	4		
Facial hair	0	1	2	3	4		
Voice							

neglected to describe the procedure that he followed to measure penis length and thickness, scrotum, and voice. But, he did report that he determined the boys' testicular size by palpation. When rating pubic hair, axillary hair, chest hair, and facial hair, Kubitschek used 0 to represent complete absence and 4 to represent adult growth. After he assessed all of the boys, he returned to their sexual development assessment tables for each annual age level, and he classified the boys' sexual development as average, retarded, or advanced (Table 2.47). For clarification, the author has placed the proportion value in parentheses beside the observed value. Although Kubitschek announced that he expected the Negro boys to sexually mature earlier than the

Table 2.47

The observed frequency of boys at each age level that were retarded, average or late maturing (Kubitschek, 1932)

Age Groups	N	Kubitschek's Classifications Retarded	Average	Advanced
9	50	8(16%)	36(72%)	6(12%)
10	50	8(16%)	33(66%)	9(18%)
11	65	14(21.5%)	38(58.5%)	13(20%)
12	80	26(32.5%)	40(50%)	14(17.5%)
13	100	30(30%)	52(52%)	18(18%)
14	110	40(36.4%)	47(42.7%)	23(20.9%)
15	105	36(34.3%)	46(43.8%)	23(21.9%)
16	100	22(22%)	51(51%)	27(27%)
17	50	12(24%)	25(50%)	13(26%)
18	20	5(25%)	12(60%)	3(15%)

white boys, he found no difference between the Negro and white boys in onset and rate of sexual maturation.

When he discussed the results of his sexual development assessments, Kubitschek divided his sample into boys 12 years old and under and boys 13 years old and older. Whereas he classified the younger boys solely on the size of their penis, scrotum, and testes, he classified the older boys chiefly on the development of their secondary sexual characteristics (Table 2.48). Kubitschek concluded that the variation in the degree of sexual development in the annual age levels showed that chronological age indexed sexual development very poorly.

In 1935, George Williams College Professor Henley S. Dimock reported on his longitudinal survey of 200 adolescent males (Dimock, 1935). As his pubescence assessment procedure, Dimock used the "Crampton criteria." Dimock credited W. B. West's Masters thesis at the International YMCA College in Springfield, Massachusetts, for

Table 2.48

The description of the sexual development changes at each age level (Kubitschek, 1937)

Age	Description
9	Small peanut size testicles mainly. A few testicles slightly larger. The degree of variation betweens boys was very small. Penis length was 3 to 4 cms.
10	Very little increase in testes size. The degree of variation between boys was very slight.
11	Same as the 10 year old group.
12	Testicles begin a fairly rapid size increase to the size of a small almond. The degree of variation between boys was considerable, as many boys still had peanut size testicles. Penis length was 3 to 7 cms. Pubic hair was frequently encountered (25% of the cases).
13	The majority of the boys initiated marked sexual development. The outstanding sexual development features of this age group were pubic hair growth, voice deepening, and masculine body configuration. Also, these boys displayed a marked testes size increase. No axillary hair present. The degree of testes size variation between boys was extensive.
14	Marked sexual development has continued. The degree in sexual development variation between boys was still extensive. Pubic hair was moderately long and dense and covered a 3 cm. diameter area above and lateral to the base of the penis. Some axillary hair present in the average and advanced boys. Husky and uneven vocal quality in some boys. The degree of testicular development variation was greatest for this age level than any other age level. Also, the degree of penis size development variation was equally as great for this age level.
15	The sexual development rate has decreased. However, the degree of sexual development variation is still very marked. The secondary sexual characteristics develop more than the primary sexual chacteristics. Long, dense, pubic hairs extend laterally and cephally. Thin, but definite, axillary hair present in almost all of the boys. Some boys have mustache hair and sideburn hair. The Average boys had large almond or small walnut sized testicles.

Table 2.48 (continued)

Age	Description
16	The sexual development has slowed considerably. The outstanding sexual development feature of this age group was the continued secondary sexual characteristics development and slowed primary sexual characteristics development. Pubic hairs extended 4 to 5 cms. lateral of the penis. Moderately long, but thin axillary hair present. Definite facial hair present. No chest hair present. Testes and penis size only slightly larger than the 15 year old group.
17	Greater degree of testes and penis size uniformity. The outstanding sexual development feature of this age group was the pubic hair growth up the linea alba. Moderately long, thick axillary hairs present. Prominent facial hair present. No chest hair present, except in several of the Advanced rated boys. Penis and testes size was more uniform.

determining that the pubic hair development index more reliably estimated sexual development than indices based on either axillary or mammary hair development. For his longitudinal study, Dimock created six classifications of pubescence status and change (Table 2.49).

Table 2.49

The six pubescent classifications that Dimock developed (Dimock, 1935)

Classification	Percent	Description
P1-P1	8%	Boys he rated pre-pubescent both years
P1-P2	13%	Boys he rated pre-pubescent the first year, and, pubescent the second year
P2-P2	1%	Boys he rated pubescent both years
P1-P3	36%	Boys he rated pre-pubescent the first year, and, post-pubescent the second year
P2-P3	26%	Boys he rated pubescent the first year, and, post-pubescent the second year
P3-P3	16%	Boys he rated post-pubescent both years

With those six classifications, Dimock analyzed the boys' height and weight. For both height and weight, Dimock found the smallest to largest change in the boys who went from P1-P1, P1-P2, P2-P2, P3-P3, P1-P3, and P2-P3. During the year the boys entered post-pubescence (P2-P3), Dimock determined that they added twice as much height and weight as the boys who remained pre-pubescent (P1-P1) for the same year period. After conducting several strength measurements on the boys, Dimock concluded that the most rapid strength development occurred after the boys reached post-pubescence (P3-P3). Based on the Brace motor ability test results, Dimock reported that except for his P1-P2 classification, the boys' motor ability increased throughout adolescence, and that substantial motor coordination improvements accompanied the boys' rapid growth in height and weight. Because the P1-P2 classification failed to improve their motor coordination, he concluded that adolescent awkwardness matched the P1-P2 boys' sudden growth onset rather than the later rapid growth period.

To indicate whether or not the University of Chicago Laboratory School boys had attained puberty, University of Chicago Professor Herman G. Richey classified boys according to the age when their axillary hair appeared (Table 2.50) (Richey, 1937). Richey found that

Table 2.50

Richey's maturity group descriptions based on time of axillary hair appearance (Richey, 1937)

Group	Description
1	Boys with axillary hair present on or before 14th birthday
2	Boys with no axillary hair present on their 14th birthday, but, who had axillary hair present on their 15th birthday.
3	Boys with no axillary hair present by their 15th birthday

boys in Maturity Group 1 were heavier at all ages than those in Maturity Groups 2 and 3, and boys in Maturity Group 2 were heavier at all ages than those in Maturity Group 3.

To determine the activity of androgenic and estrogenic hormones extracted from the urine of adolescent males, Yale University Professors Ralph I. Dorfman, William Walter Greulich, and Charles I. Solomon collected the total daily urinary excretion of 18 boys aged 93 to 195 months (Dorfman, Greulich, and Solomon, 1937). By dissolving the androgenic material of the urine in olive oil, subcutaneously injecting that solution into day old white leghorn chickens, and then determining the increase in the weight of the chickens' combs, they measured the amount of androgenic activity. They measured the estrogenic activity of the urine by D'Armour's and Gustavson's adult spayed rat vaginal smear technique. Table 2.51 contains the results of those assays for the eighteen boys that they measured. They noted that Cases 5 and 14 were physically more mature than the other boys. They concluded that androgenic and estrogenic activities varied considerably for boys of the same chronological age. Therefore, as their next investigation they proposed to examine the relationship between boys' sex hormone secretions and degree of physical maturity.

Using 24 hour urine specimens from 18 prepubertal boys aged 11 years and under and 15 adolescent boys aged 12 to 16 years, Yale University Professors H. R. Catchpole, W. W. Greulich, and R. T. Sollenberger measured the urinary excretion of follicle stimulating hormone (FSH). They found that none of the 18 prepubertal boys showed any FSH, but eleven of the 15 adolescent boys showed adult amounts of

Table 2.51

The amount of androgenic and estrogenic activity found in the urine of adolescent males (Dorfman, Greulich, and Solomon, 1937)

Case Number	Chronologic Age	Androgenic Activity	Estrogenic Activity
1	93 months	1.1	9.0
2	108 months	2.0	5.0
3	132 months	2.0	15.0
4	132 months	2.0	10.6
5	158 months	15.0	48.0
6	158 months	1.2	9.0
7	158 months	4.0	12.0
8	159 months	5.3	24.0
9	160 months	8.0	13.0
10	162 months	8.0	10.0
11	171 months	7.0	--
12	176 months	16.5	38.0
13	181 months	6.0	18.0
14	184 months	32.0	67.0
15	189 months	20.0	65.0
16	190 months	8.4	24.0
17	194 months	14.0	95.0
18	195 months	19.0	68.0

FSH. They concluded that the urinary FSH of adolescent males related to their developmental status rather than their chronological age (Catchpole, Greulich, and Sollenberger, 1938).

To gain more information about the normal variation in the developmental sizes of male genitalia, New York Physicians William A. Schonfeld and Gilbert W. Beebe used a centimeter ruler, a circumference gauge, and an orchidometer to measure the penis length and circumference, and testicular volumes of males (Schonfeld and Beebe, 1942). They hypothesized that the true physiological length of the penis was its erect length. However, because they believed they could not feasibly obtain measurements on the erect penis of their subjects, they substituted a measure of the fully stretched length of the flaccid penis for the erect penis length. They measured from the mons veneris to the tip of the glans for the penis length measurement. To test the validity of their penis length measurement, they asked 150 males of different ages to outline their erect penis on cardboard for four successive mornings. By correlating their measurements of the subjects' drawings with their measurements of the fully stretched but flaccid penis length, they obtained a coefficient of .983. Based on that result, they stated that their measurement of the fully stretched but flaccid penis was a valid estimate of the erect penis length.

To obtain the penis circumference measure, they used a set of circumference gauges. After they slipped a circumference gauge over the glans, they determined whether or not that gauge would easily slide over the shaft of the fully stretched penis. Therefore, they measured the largest circumference of the fully stretched but flaccid penis. To test the validity of the penis circumference measurement, they asked the same 150 subjects to insert their erect penis through measured holes punched into cardboard for four successive mornings. The subjects reported the smallest hole through which their erect penis fit with ease. By correlating the subjects' findings with their circumference gauge measurement, they obtained a coefficient of .965. Based on that result, they stated that with the aid of the regression equation $Y = .1362 + 1.316X$, where Y is the predicted penis circumference and X is the observed circumference of the stretched but flaccid penis circumference, their measurement of the flaccid penis circumference was a valid estimate of the erect penis circumference.

To obtain the testicular volume, they simultaneously palpated their standard set of graded orchidometer models and the testes. They recorded the volume of the orchidometer model that they determined was the same size as the boys' testes. To test the validity of the testicular volume measurements, they used this orchidometer to estimate the volume of 400 testicular clay models and to estimate the volume of 35 postmortem excised human testes. They found that they frequently overestimated the volumes of the smaller testicular clay models and postmortem testes and that they frequently underestimated the volumes of the larger testicular clay models and postmortem testes. However, because they determined that their average error was ten percent or less, they stated that their orchidometer model measurements validly estimated the boys' testicular volume (Table 2.52).

They concluded that male genitalia growth proceeded so rapidly that for any pubescent age the genitalia measurements skewed towards the higher values, and that pubescent boys' testicular volume measurements were bi-modally distributed. Because the testicular volume increased so rapidly at pubescence, Schonfeld and Beebe hypothesized that researchers could use this rapid testicular size increase as an indicator of the onset of male pubescence. Because the pubescent rate of genitalia growth far outstripped the rate of growth in height and weight, they concluded that genitalia growth was a phenomenon of pubescence. After their subjects reached ten years of age, Schonfeld and Beebe argued that chronological age no longer reliably estimated genital development.

Table 2.52

Genitalia measurements on males from birth to adult
(Schonfeld and Beebe, 1942)

Age (months)	N	Penis Length	Relaxed Penis Circumference	Erect Penis Circumference	Testicular Volume	
0-5	125	3.75cm	3.63cm	4.91cm	.52cc	
6-11	51	4.04	3.70	5.01	.72	
12-23	50	4.59	3.89	5.26	.74	
34-35	50	5.04	4.06	5.48	.83	
36-47	50	5.43	4.06	5.47	.80	
48-59	50	5.75	4.10	5.53	.82	
60-71	50	6.00	4.19	5.65	.79	
72-83	50	6.02	4.27	5.76	.75	
84-95	50	6.17	4.46	6.01	.82	
96-107	50	6.21	4.23	5.70	.83	
108-119	50	6.32	4.45	5.99	.94	
120-131	82	6.20	4.57	6.15	.95	
132-143	91	6.56	4.74	6.37	1.48	Pubescent Age
144-155	93	7.13	5.05	6.78	2.22	
156-167	104	8.73	5.79	7.76	4.86	
168-179	121	9.77	6.88	9.19	7.72	
180-191	101	11.81	7.62	10.16	11.76	
192-203	76	12.50	7.99	10.65	12.61	
204-215	61	13.26	8.43	11.23	15.12	
216-239	71	13.11	8.61	11.47	16.06	
240-300	54	13.02	8.55	11.39	16.47	

At the Adolescent Study Unit and Institute of Human Relations of Yale University's School of Medicine, Professors William Walter Greulich, Ralph I. Dorfman, and Hubert R. Catchpole and Physicians Charles I. Solomon and Charles S. Culotta studied selected changes in somatic growth and endocrine function that occurred during the male adolescence period (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942). Professor Dorfman reported on the urinary excretion of estrogens and androgens, Professor Catchpole reported on the urine levels of pituitary gonadotrophic hormone, and Professor Greulich reported on skeletal and other somatic changes. Physicians Solomon

and
Sev
con
the
Cen
Ph
di
ch
ob
th

Pr
se
an
da
P
3
a
c
c
H
c
c
t
c
a

and Culotta assisted the three professors on each of their reports. Seven hundred fifty-seven white males ranging in age from 10-15 years comprised their sample. One hundred ninety-four of the boys were in the Connecticut School for Boys, 32 resided in the New Haven Community Center orphanage, 476 attended the large, expensive Andover, Massachusetts, Phillips Academy private school, and 55 of the boys lived at home. To direct special attention to the nature and sequence of external bodily changes associated with adolescent males' sexual maturation, they obtained nude photographs of the boys. In addition to the photographs, the investigators described each boy's external sexual development.

By examining the photographs and the written descriptions, Professor Greulich outlined the development of the boys' primary and secondary sexual characteristics. Greulich reported on 1) the growth and development of the boys' external genitalia, 2) the appearance and development of the boys' head hair, facial hair, pubic hair, scrotal, perineal, and cirumanal hair, axillary hair, and limb and trunk hair, 3) the activity of the boys' axillary sweat glands, 4) the appearance and activity of the boys' mammary gland, and 5) the boys' voice changes. Without providing ages, Greulich listed the appearance order of the external sexual characteristic changes in his boys (Table 2.53). He cautioned that the order of appearance designated only the onset of the development of those sex characteristics, but that each sex characteristic developed at its own rate thereafter. Greulich stated that because of the adolescent boy's great variability in sexual development, specialists in child growth should consider chronological age only as a measure of how long the boys had lived and not an index

Table 2.53

The order of appearance of the external sexual characteristics in males (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942)

Order	Description
1	Accelerated testes growth followed by accelerated penis growth.
2	Long, downy pubic hairs appear at the base of the penis
3	Long, coarse, straight, pigmented pubic hairs appear at the base of the penis.
4	Axillary perspiration markedly increases.
5	The downy vellus at the corners of the upper lip becomes longer, coarser, and darker.
6	Long, coarse, downy hairs appear on proximal third of their forearm's extensor surface and on the distal fourth of their arm's lateral surface.
7	Coarse, moderately pigmented hair appears on the distal half of their legs' extensor surfaces and on the distal third of the legs' flexor surfaces.
8	Long, downy, hairs appear at their sideburns areas.
9	Circumanal hairs appear.
10	Moderate to dense pubic hair appear and covers the pubic triangle region.
11	Short, fine, pigmented hairs appear in their axilla.
12	Their voice deepens.
13	When present, their subareolar masses attain their maximum sizes.
14	Their pubic hairs spread totheir thighs' medial surfaces and up the linea alba.
15	Their penis and testes attain their full adult dimensions.
16	A few terminal hairs appear on their chins and sideburns areas, and, their upper lip hairs become coarses and darker.
17	A few terminal hairs appear around their areolae and on their sternum.
18	Their head hairlines differentiate toward adult configuration.
19	Adult hair quality and quantity present on their forearms, arms, legs, and thighs.

of the amount of progress that they had made toward their adult maturity. Therefore, he suggested a procedure to classify adolescent boys according to sexual development. Greulich proposed five categories that encompassed the successive developmental stages of adolescent boys' external primary and secondary sexual characteristics (Table 2.54). From his categories,

Table 2.54

Descriptions of the sexual characteristic changes during adolescence in males (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942)

Group	Description
1	a. Early childhood penis, testes, and scrotum size and configuration. b. Pubic hair no better developed than abdominal hair c. Early childhood facial hair d. No definitive circumanal, perineal, or axillary hair
2	a. Definitely larger testes b. Some penis growth c. Conspicuous, long, lightly pigmented, downy hairs at the base of the penis d. Early childhood facial hair e. No definitive circumanal, perineal, or axillary hair
3	a. Testes somewhat larger b. Penis length has increased c. Pubic hairs have increased in length, coarseness, and amount, and, occasional long, coarse, straight pigmented hair intersperse the pubic region d. Half of the boys have short, lightly pigmented, upper lip hair and some similar sideburn hair e. A few boys had some short, coarse, pigmented circum-anal hair f. A few boys had some short, fine, lightly pigmented axillary hair
4	a. Testes somewhat to considerably larger b. Penis diameter has increased c. Adult type pubic hair covers the pubic triangle d. Small to moderate amount of short, lightly pigmented coarse upper lip hairs, and a similar amount of long, fine, unpigmented cheek hair e. A few moderately long, coarse, pigmented, chin hair f. No throat or sub-mandibular hair g. Well-differentiated circumanal hair h. Axillary hair is less dense and undifferentiated than adult axillary hair
5	a. Adult proportion testes b. Adult configuration and proportion penis, may appear disproportionately large c. Adult quality and quantity pubic hair and extends toward the thighs and linea alba d. Adolescent mustache e. Relatively long, moderately coarse, lightly pigmented sideburn hairs f. Moderate growth of chin hair g. Well-developed circumanal and axillary hair

researchers could describe the adolescent boys' sexual development status.

Greulich cited research on both animals and humans that indicated skeletal maturation and sexual maturation related closely to each other. The research clearly established that normal skeletal maturation did not progress without adequately functioning gonads. As evidence, he reported that hypogonadal or castrated adolescent males displayed delayed epiphysial fusion. Greulich compared adolescent boys' skeletal maturation with their primary and secondary sexual characteristics development. He X-rayed the hand-wrists and elbows of the 476 private school boys in the sample. Using Todd's 1937 Maturation Atlas, Brush Foundation worker Clara Buehl assessed the boys' skeletal age (Table 2.55). Because he had insufficient sample sizes in Maturity Groups 1 and 2, the mean skeletal age values fluctuated erratically. However, beginning with Maturity Group 3, the mean skeletal ages increased systematically. The mean skeletal age from Maturity Group 3 to Maturity Group 3-4 increased 3.5 months, from Maturity Group 3-4 to Maturity Group 4 5.7 months, from Maturity Group 4 to Maturity Group 4-5 7.7 months, and from Maturity Group 4-5 to Maturity Group 5 10.7 months. Greulich noted that within a single maturity group, the older boys tended to be more skeletally mature than the younger boys. However, the older boys were less skeletally mature than their chronological ages indicated. Within the same maturity group, he noted that the younger boys skeletally matured earlier than their chronological ages indicated.

Table 2.55

The number and mean skeletal age values of 476 adolescent males (Greulich, Dorfman, Catchpole, Solomon and Culotta, 1942)

Groups		Chronological Age (months) 144.0	156.0	168.0	180.0	192.0	204	N	X
1	N	4	0	0	1	0	0	5	156.4
	X	154.2	XX	X	165.0	XX	XX		
1-2	N	0	0	4	0	0	0	4	157.2
	X	XX	XX	157.2	XX	XX	XX		
2	N	0	1	6	1	0	0	8	166.5
	X	XX	164.0	167.5	163.0	XX	XX		
2-3	N	1	1	1	1	0	0	4	162.8
	X	165.0	165.0	148.0	173.0	XX	XX		
3	N	0	8	21	7	0	1	37	166.8
	X	XX	161.8	167.0	170.8	XX	179.0		
3-4	N	0	3	14	10	0	0	27	170.6
	X	XX	168.7	168.9	173.4	XX	XX		
4	N	1	7	41	35	5	0	89	176.3
	X	163.0	167.3	175.3	179.1	180.4	XX		
4-5	N	1	10	42	43	13	3	112	184.7
	X	176.0	180.4	181.1	186.6	191.1	197.7		
5	N	1	4	45	89	46	5	190	195.4
	X	181.0	180.2	189.8	196.2	199.2	209.0		
TOTAL	N	4	38	174	187	64	9		
	X	171.2	169.5	178.1	188.2	196.1	201.9		

For their determinations of the estrogenic and androgenic levels in the urine samples from 89 boys aged 4 days to 22 years with most of the boys aged 93 months to 206 months, they used vaginal responses of adult spayed female mice and comb growth response of day-old chicks, respectively. Professor Dorfman provided the androgenic material level (I.U.) and estrogenic material level (I.U.) by Greulich's Maturity Groups, and estrogenic material level (I.U.) by chronological age groups (Tables 2.56, 2.57, and 2.58).

For their determinations of the amount of pituitary gonadotrophic hormone in the adolescent boys' urine, they used the same urine samples that Dorfman used for his estrogenic and androgenic

Table 2.56

The androgenic material level (I.U.) classified by Maturity Groups (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942)

Descriptive Data	Maturity Groups				
	1	2	3	4	5
N	11	19	22	20	20
$\bar{X}$	4.09	6.63	10.04	19.70	28.80
S.D.	2.50	3.36	6.36	10.25	15.30
Range	1-8	2-16	2-26	7-53	9-57

Table 2.57

The estrogenic material level (I.U.) classified by Maturity Groups (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942)

Descriptive Data	Maturity Groups				
	1	2	3	4	5
N	6	11	16	17	20
$\bar{X}$	7.00	11.36	16.69	37.76	45.85
S.D.	3.06	7.56	11.49	15.44	21.80
Range	2-10	3-30	3-45	15-75	17-95

Table 2.58

The estrogenic material level (I.U.) classified by chronological age groups (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942)

Descriptive Data	Chronological Age Groups				
	99-155	156-167	168-179	180-191	192-204
N	9	15	12	19	15
$\bar{X}$	16.56	13.87	25.58	39.68	38.33
S.D.	13.52	13.03	18.13	21.66	20.96
Range	2-48	3-50	5-75	3-75	8-95

material study. To bio-assay for pituitary gonadotrophic hormone, they used 21-23 day old female mice. Professor Catchpole compared gonadotrophin classifications with chronological age groups and Greulich's five Maturity Groups (Tables 2.59 and 2.60). Catchpole concluded that as the adolescent males' sexual development advanced, their urine

Table 2.59

A comparison of gonadotrophin classifications with chronological age (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942)

Chronological Age (months)	Gonadotrophin Classifications 3	3-7	7
109-144	21	0	0
145-162	8	3	0
163-180	12	8	3
181-198	10	10	10

Table 2.60

A comparison of gonadotrophin classifications with the Maturity Groups (Greulich, Dorfman, Catchpole, Solomon, and Culotta, 1942)

Maturity Groups	Gonadotrophin Classifications 3	3-7	7
1	5	0	0
2	7	1	0
3	4	3	0
4	6	7	4
5	2	6	8

gonadotrophin levels generally tended to increase from early childhood's undetectable levels to levels characteristic of adults.

To determine the chronological age range in the onset of male pubescence, Physician William A. Schonfeld measured the penis length and circumference, and testes volume of 1500 randomly selected males aged 1 day to 25 years (Schonfeld, 1943a). And, from the adolescent males within that sample, he assessed the degree of maturation in secondary sexual characteristics. To subdivide those boys into sexual maturation groups, Schonfeld incorporated Crampton's Pubic Hair Index and created six sexual maturation stages (Table 2.61). In addition to noting the development in boys' secondary sex characteristics, Schonfeld determined the statistical range of their genitalia growth

Table 2.61

Schonfeld's six sexual maturation stages (Schonfeld, 1943a)

Stage	Description
1	Prepubescent boys, no sexual development present
2	Active growth of the penis and testes, but no pubic hair present
3	Crampton's Pubic Hair Index Stage 1 = prepubescence
4	Crampton's Pubic Hair Index Stage 2 = pubescence
5	Crampton's Pubic Hair Index Stage 3 = post-pubescence
6	Adult sexual maturation

measures. Based on data that he displayed in a graph showing the relationship between testes volume and age, Schonfeld concluded that male pubescence began at any age between 10 and 16 years.

Much of Schonfeld's report discussed the theory and practice of endocrine therapy for hypogonadal boys. Basically, his treatise was that physicians must be certain that they have properly diagnosed the problem before they treat boys for hypogonadism.

Without an accompanying written description, Schonfeld provided a schematic table of the stages of sexual development. In that schematic table, Schonfeld had drawings of hairline, facial hair, chin, larynx, breasts, axillary hair, body configuration, body hair, pubic hair and prostate development. Also, he provided the size range in the six sexual maturation stages for penis length, penis circumference, and testes volume. By extracting the data for penis length and circumference, and testes volume from that schematic table, the author created Table 2.62. Inspection of this table demonstrates the tremendous overlap in the size ranges for each of Schonfeld's stages. For example, a boy with a penis length of 8 cm. could belong in any of Schonfeld's first four stages.

Table 2.62

The size range within the six sexual maturation stages for penis length, penis circumference, and testes volume (Schonfeld, 1943a)

	Schonfeld's Six Sexual Maturation Stages					
	1	2	3	4	5	6
Penis length (mm.)	3-8	4.5-9	4.5-12	8-15	9-15	10.5-18
Penis circum. (cm.)	3-5	4-6	4-8	4.5-10	6-10	6-10.5
Testes volume (cc.)	.3-1.5	1.75-6	1.75-13	2-20	6-20	8-25

Without an accompanying written explanation, Schonfeld provided a table that gave the frequency of the boys in annual age increments that he rated in his six sexual maturation stages (Table 2.63).

Table 2.63

The proportion of boys at each age in the six sexual maturation stages (Schonfeld, 1943a)

Age	Schonfeld's Sexual Maturation Stages					
(Years)	1	2	3	4	5	6
9	100.0	0.0	0.0	0.0	0.0	0.0
10	94.0	0.0	0.0	0.0	0.0	0.0
11	76.0	12.0	12.0	0.0	0.0	0.0
12	44.0	14.0	32.0	10.0	0.0	0.0
13	15.0	18.0	38.0	21.0	8.0	0.0
14	6.0	15.0	26.0	26.0	27.0	0.0
15	0.0	2.0	16.0	22.0	53.0	7.0
16	0.0	1.0	9.0	11.0	59.0	20.0
17	0.0	0.0	3.0	7.0	39.0	51.0
18	0.0	0.0	0.0	7.0	30.0	63.0
19	0.0	0.0	0.0	0.0	26.0	74.0
20-21	0.0	0.0	0.0	0.0	17.0	83.0
22-25	0.0	0.0	0.0	0.0	0.0	100.0

Pubescence began as early as age 10 for four percent of the boys and by age 14 ninety-four percent of the boys had started pubescence. Also, seven percent of the boys completed their sexual maturation by age 15 and only fifty-one percent of the boys had completed their sexual maturation by age 17.

To evaluate the variation in sexual maturation of normal boys, William A. Schonfeld also assessed the genital status of his subjects (Schonfeld, 1943b). Schonfeld excluded any boys with known or suspected endocrinopathy or other associated abnormality in genitalia development. As primary sexual characteristics, Schonfeld listed the testes, their epididymides, the penis, the scrotum, and the prostate. Schonfeld advised that when considering testes growth, in addition to testicular volume increases, researchers should consider the relationship of the testes to the epididymides. He found that while at birth the testes and their epididymides are about equal in size, during early pubescence because adolescent males' testes grow at a more rapid rate, their epididymides are only about one-quarter of their testes size, and in adults the epididymides are about one-ninth the size of their testes. Further, during the adolescent males' rapid growth period, epididymides bind more closely to their testes. Regarding the relative size of the boys' right and left testicles, Schonfeld determined that in 23 percent of the boys the size of the right testicle exceeded the left, in 22 percent of the boys the size of the left testicle exceeded the right, and the remaining 55 percent of the boys had testicles equal in size. In his discussion of the scrotal changes associated with the onset of male puberty, Schonfeld observed that to accommodate the testes enlargement, the scrotal sac distended at its distal end while its proximal end became relatively narrower.

In his discussion of the prostate gland changes associated with the onset of male puberty, Schonfeld observed that as the result of its five undifferentiated lobes enlarging, the prostate's monolobar organ progressively increased in size. At puberty, he noted that the two lateral lobes enlarged relatively greater than the other prostate lobes, thereby creating a midline groove. The posterior lobe capped that midline groove, the anterior lobe shrunk into insignificance, and the middle lobe enlarged similar to the lateral lobes. After the boys reached that degree of prostate differentiation, Schonfeld reported that physicians could massage the prostate and cause the prostate to secrete its contents. Schonfeld reported that he had massaged such prostates, and, that he was studying the microscopic constituents of those secretions.

As secondary sexual characteristics, Schonfeld listed hair development, breast changes, and voice or larynx growth. Despite the wide variation in the degree of hair development among individuals, families, and races, he considered hair development an important evaluative factor in assessing the sexual development of adolescent males. In his discussion of pubic hair development, Schonfeld described four stages (Table 2.64).

Table 2.64

Pubic hair stages (Schonfeld, 1943b)

Stage	Description
1	Darkening of the villous hair at the base of the penis
2	Extension of the pubic hair laterally to the inguinal region
3	Upward growth of the pubic hair to cover the mons pubis triangle
4	Extension of the pubic hair to the umbilicus

Schonfeld combined the hair growth in the axillary, on the upper lip, on the face, body, and perineal areas into one discussion.

He said that adolescent boys' axillary hair developed after their pubic hair completed stage two, mustache hair developed concurrent with axillary hair, chin and sideburn hair developed shortly after axillary and mustache hair, and body and perineal hair developed after pubic, axillary, mustache, chin, and sideburn hair, but well before the facial hair. Schonfeld found forehead hairline change one of the last secondary sexual characteristics to occur. He described that immature boys' uninterrupted, bowlike forehead hairline changed to the adult double lateral wedge-shaped indententation, interrupted forehead hairline.

Regarding adolescent males' breast development, Schonfeld reported that unilateral or bilateral subareolar nodules occurred and receded at varying pubescent stages and the degree of adolescent male breast development varied from small, painless, subareolar nodules to pronounced, painful subareolar nodules that secreted a fluid.

Regarding the larynx development of adolescent males, Schonfeld determined that in addition to rapid growth, the boys' larynx changed its form, and the enlarged thyroid cartilage's ventrodorsal diameters created the adolescent males' laryngeal prominence. Concurrent with laryngeal prominence, the vocal folds lengthen and thicken, thereby dropping the pitch in the voice by one octave.

To determine the mean height and weight differences of adolescent males when grouped by sexual maturity, Professor Richard W. B. Ellis of the University of Edinburgh Department of Child Life and Health reviewed descriptions of physical growth and primary and secondary sexual characteristics that trained personnel at two residential schools made annually on 208 boys aged 11-16 years (Ellis, 1946). At both schools, the medical personnel recorded presence and extent of pubic, axillary, facial, and body hair, and they described the boys' genitalia development. Also, by paying special attention to acne and sebaceous and apocrine sweat gland development, they assessed the boys' skin appearance.

Without knowledge of Greulich and his associates' report, from those primary and secondary sexual characteristics development descriptions, Ellis created pubic hair and axillary hair sexual maturation stages (Tables 2.65 and 2.66). Ellis found facial hair difficult

Table 2.65

Pubic hair stages (Ellis, 1946)

Stage	Description
PO	Early childhood fine body hair or slightly longer, non-pigmented vellus present
P1	One or more coarse, pigmented hairs appear at the base or sides of the penis
P2	Pigmented hairs extend across the root of the penis
P3	Pigmented hairs cover approximately half of the pubic triangle
P4	Pigmented hairs cover the entire pubic triangle

Table 2.66

Axillary hair stages (Ellis, 1946)

Stage	Description
A0	Complete absence of hair
A1	One or more pigmented hairs present
A2	One centimeter long, pigmented, hair present and they cover a two centimeter diameter area
A3	Profuse axillary hair present

to asses, and therefore he rated facial hair as either present or absent.

Regarding genitalia development, Ellis described the growth of the penis, testes, and scrotum separately. First the penis lengthened considerably, then because of corpora cavernosa development, the penis thickened considerably. Because the testes of adolescent males grew more rapidly than their epididymides, the relative testes-epididymides proportion differed markedly from childhood, and that concurrent with the spurt in testicular size, the testes softened. Not only did the scrotum become pendulous with the enlarged testicles, but it also changed from being wider at the proximal end to being wider at the distal end. Whereas the testicles and scrotum developed synchronously, their testicles and penis did not necessarily develop in synchrony. Rather than create specific stages for those three components of genitalia, Ellis simply rated the boys' genitals as prepubescent, pubescent, and adolescent.

Because he found that observations of facial acne failed to develop sequentially, Ellis did not use the skin observations as criteria for sexual maturation. He considered the observations on the development of the apocrine and sebaceous glands as confirmatory and not diagnostic. Using those criteria of secondary sexual characteristics and genitalia development as his basis, Ellis created three classes of sexual maturation in adolescent males (Table 2.67).

To verify the reliability of his adolescent male sexual maturation classes, Ellis rated the 132 boys for a second time one hundred days later. Because in the second rating he did not place any boys

Table 2.67

Three sexual maturation classes for adolescent males (Ellis, 1946)

Class	Description
Prepubescent	No development present
Pubescent	Pubic hair present and/or pubescent genital development present
Adolescent	Profuse pubic hair present and considerable growth in penis length and girth and testes volume

in a class lower than he had in the first rating, Ellis regarded his classification procedure to be reliable. Ellis gave the proportion of the boys aged 11 to 16 that he rated as prepubescent, pubescent, and adolescent (Table 2.68).

Table 2.68

The proportions of the boys at each age in the three sexual maturation classes (Ellis, 1946)

Age		Ellis' Sexual Maturation Classes		
(Years)	N	Prepubescent	Pubescent	Adolescent
11 to 12	26	88.5	11.5	0.0
12 to 13	39	64.1	35.9	0.0
13 to 14	73	42.5	46.6	11.0
14 to 15	44	13.6	25.0	61.4
15 to 16	26	0.0	30.8	69.2

Although he used cross-sectional data, Ellis concluded that pubic hair appeared when genital development advanced to pubescent and that axillary hair appeared when genital development advanced to adolescent. From Ellis' discussion regarding the extent of genitalia, pubic hair and axillary hair development of boys that he rated as pubescent and adolescent, the author created Table 2.69.

Beginning when the boys were nine years old, for twelve successive years Fels Research Institute for the Study of Human Development

Table 2.69

The proportion of the boys that Ellis rated as pubescent and adolescent with various genitalia, pubic hair, and axillary hair ratings (Ellis, 1946)

Ellis' Classification	Percent	Genitalia Rating	Pubic Hair Rating	Axillary Hair Rating
	10%	Prepubescent	P1 to P3	A0
	23%	Pubescent	P0	A0
Pubescent	52%	Pubescent	P1 to P3	A0
	12%	Pubescent	P1 to P3	A1 or A2
	3%	Prepubescent	P0	A3
	48%*	Adolescent	P2 to P5	A0
Adolescent	48%*	Adolescent	P2 to P5	A1 or A2
	4%*	Adolescent	P1 to P3	A3

(*Those percentages are estimates)

workers semi-annually photographed, measured, X-rayed, and wrote sexual characteristic development descriptions on 59 boys (Reynolds and Wines, 1948). Professor Earle L. Reynolds and Janet V. Wines constructed a five-point genitalia maturation scale (Table 2.70).

Table 2.70

Five-point genitalia maturation scale (Reynolds and Wines, 1948)

Rating	Description
1	Infantile development
2	Scrotum enlarges, reddens, and changes texture
3	Penis enlarges and sculptures
4	Pronounced penis sculpturing and darkening
5	Adult reddish brown color, with lose penile skin and loss of sharp sculpturing

They provided photographs that coincided with each of their five genitalia stages. Reynolds and Wines calculated the proportion of the boys that they rated in each of their five genitalia stages (Table 2.71). The average boy required less than six years to complete

Table 2.71

The proportion of the boys at each age in the five genitalia stages (Reynolds and Wines, 1948)

Ages (Years)	N	Reynolds-Wines Genitalia Maturation Stages 1	2	3	4	5
9.0	59	100.0	0.0	0.0	0.0	0.0
9.5	59	97.0	3.0	0.0	0.0	0.0
10.0	57	95.0	5.0	0.0	0.0	0.0
10.5	52	83.0	17.0	0.0	0.0	0.0
11.0	49	65.0	33.0	0.0	0.0	0.0
11.5	47	36.0	55.0	2.0	4.0	0.0
12.0	43	23.0	63.0	4.0	5.0	0.0
12.5	43	9.0	42.0	9.0	7.0	0.0
13.0	39	3.0	18.0	42.0	38.0	0.0
13.5	39	0.0	13.0	41.0	64.0	0.0
14.0	37	0.0	8.0	23.0	87.0	0.0
14.5	35	0.0	3.0	5.0	94.0	0.0
15.0	34	0.0	0.0	3.0	97	0.0
15.5	31	0.0	0.0	3.0	97	3.0
16.0	27	0.0	0.0	0.0	96	4.0
16.5	26	0.0	0.0	0.0	88	12.0
17.0	22	0.0	0.0	0.0	77	23.0
17.5	16	0.0	0.0	0.0	56	44.0
18.0	16	0.0	0.0	0.0	25	75.0

his genitalia maturation. From those proportions, the author constructed Figure 2.1.

Apart from their genitalia ratings, Reynolds and Wines rated the boys' penis size in stages from one through seven. They provided photographs and gave the proportion of the boys that they rated in each of their seven penis size stages (Table 2.72).

Reynolds and Wines provided five point pubic hair stages (Table 2.73). After rating the pubic hair development of their sample, they calculated the proportion of those boys that they rated in each of their five pubic hair stages (Table 2.74), and the author used those proportions to construct Figure 2.2.

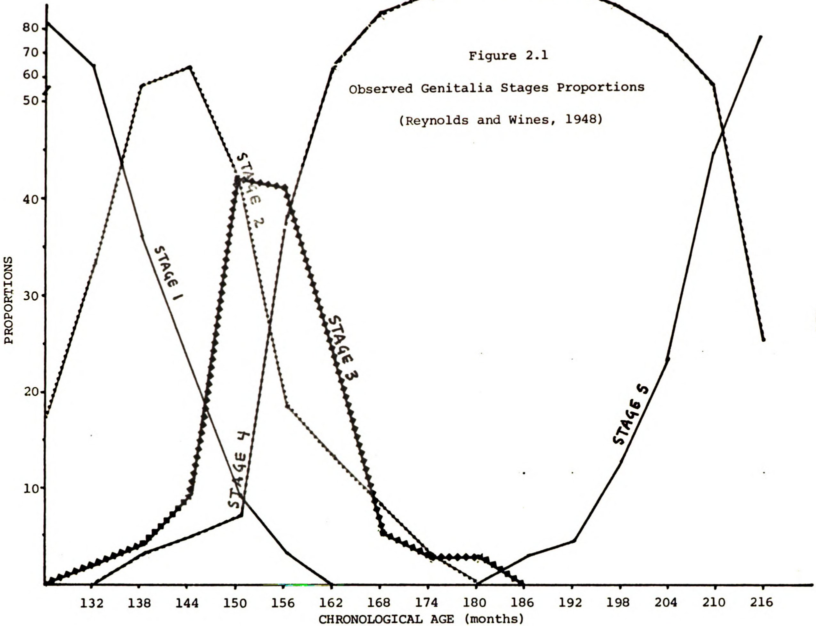


Figure 2.1

Observed Genitalia Stages Proportions

(Reynolds and Wines, 1948)

Table 2.72

The proportion of the boys at each age in the seven penis size stages (Reynolds and Wines, 1948)

Ages (Years)	N	Reynolds-Wines Penis Size Stages						
		1	2	3	4	5	6	7
10.0	58	28.0	53.0	14.0	5.0	0.0	0.0	0.0
10.5	54	26.0	54.0	15.0	5.0	0.0	0.0	0.0
11.0	50	24.0	54.0	16.0	6.0	0.0	0.0	0.0
11.5	48	25.0	48.0	17.0	8.0	2.0	0.0	0.0
12.0	43	16.0	56.0	16.0	7.0	0.0	5.0	0.0
12.5	42	12.0	36.0	40.0	5.0	2.0	2.0	2.0
13.0	39	8.0	10.0	36.0	26.0	15.0	3.0	3.0
13.5	39	3.0	13.0	8.0	36.0	20.0	15.0	5.0
14.0	37	3.0	8.0	5.0	11.0	35.0	22.0	16.0
14.5	33	0.0	3.0	3.0	18.0	27.0	33.0	15.0
15.0	31	0.0	0.0	3.0	16.0	23.0	42.0	16.0
15.5	30	0.0	0.0	0.0	20.0	20.0	40.0	20.0
16.0	25	0.0	0.0	0.0	16.0	20.0	48.0	16.0
16.5	26	0.0	0.0	0.0	15.0	23.0	42.0	19.0
17.0	22	0.0	0.0	0.0	14.0	36.0	41.0	9.0
17.5	17	0.0	0.0	0.0	18.0	35.0	41.0	6.0
18.0	15	0.0	0.0	0.0	27.0	40.0	27.0	7.0

Table 2.73

Pubic hair stages (Reynolds and Wines, 1948)

Stage	Description
1	No hair
2	Pigmented, straight, sparse pubic hair appears at the base of the penis
3	Pubic hair has curled slightly, spread slightly, and darkened
4	Pubic hair has curled, increased moderately and spread moderately, but not to the thighs
5	Adult pubic hair covers the pubic triangle

Table 2.74

The proportion of the boys at each age in the five pubic hair stages (Reynolds and Wines, 1948)

		Reynolds-Wines Pubic Hair Maturation Stages				
Ages	N	1	2	3	4	5
8.5	59	100.0	0.0	0.0	0.0	0.0
9.0	59	98.0	2.0	0.0	0.0	0.0
9.5	59	98.0	2.0	0.0	0.0	0.0
10.0	59	98.0	2.0	0.0	0.0	0.0
10.5	54	91.0	9.0	0.0	0.0	0.0
11.0	50	86.0	14.0	0.0	0.0	0.0
11.5	48	69.0	29.0	2.0	0.0	0.0
12.0	45	60.0	31.0	7.0	2.0	0.0
12.5	42	48.0	43.0	7.0	2.0	0.0
13.0	40	28.0	38.0	20.0	12.0	2.0
13.5	39	10.0	33.0	20.0	33.0	3.0
14.0	37	3.0	14.0	22.0	59.0	3.0
14.5	34	3.0	3.0	9.0	74.0	12.0
15.0	33	3.0	0.0	3.0	76.0	18.0
15.5	30	0.0	0.0	0.0	73.0	27.0
16.0	25	0.0	0.0	0.0	56.0	44.0
16.5	26	0.0	0.0	0.0	46.0	54.0
17.0	22	0.0	0.0	0.0	32.0	68.0
17.5	16	0.0	0.0	0.0	12.0	88.0
18.0	15	0.0	0.0	0.0	0.0	100.0

Reynolds and Wines noted that after the first appearance of pubic hair, the boys reached the adult stage within four years. By combining the findings on genitalia and pubic hair development, they determined the most common sequence of appearance in genitalia and pubic hair stages (Table 2.75). Although this order occurred most frequently, nevertheless forty-one percent of the boys reached Genitalia Stage 3 prior to Pubic Hair Stage 2, thirty-two percent reached Genitalia Stage 4 prior to Pubic Hair Stage 3, and seventy percent of their boys reached Genitalia Stage 2 prior to reaching Pubic Hair Stage 2. However, all boys in their sample reached Pubic Hair Stage 5 prior to reaching Genitalia Stage 5. Therefore, they

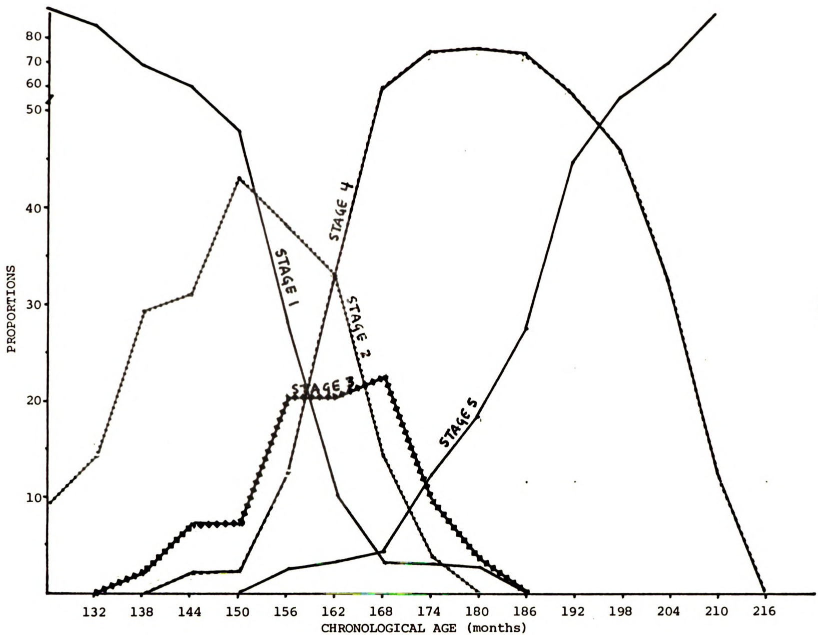


Figure 2.2 Observed pubic hair proportions (Reynolds and Wines, 1948)

Table 2.75

The most common order of appearance of the genitalia and pubic hair stages (Reynolds and Wines, 1948)

Order	Genitalia or Pubic Hair Stage
1	Genitalia Stage 2
2	Pubic Hair Stage 2
3	Genitalia Stage 3
4	Pubic Hair Stage 3 and Genitalia Stage 3
5	Pubic Hair Stage 4
6	Pubic Hair Stage 5
7	Genitalia Stage 5

concluded that the genitalia maturation of adolescent males started earlier, but finished later than pubic hair maturation.

To provide a picture of the sequence of changes associated with puberty, and to examine the variability in the onset and duration of each sequential phase, Hamilton Hogben, J. A. H. Waterhouse, and Lancelot Hogben examined 642 England secondary school boys from the Borough of Tottenham (Hogben, Waterhouse, and Hogben, 1948). To estimate the boys' sexual development, they assessed their pubic, axillary, chin and mustache hair growth, pubic hair contour, testes position, voice quality, seminal activity and urine albumin content. They recorded the development of pubic, axillary, chin, and mustache hair development as: 1) absent, 2) short, or 3) copious, and pubic hair contour as: 1) apical, 2) intermediate, or 3) horizontal. For testes position or inguinal canal closure, if the boys failed to voluntarily retract the testicles into the inguinal canal, then the examiner gently pushed the testicles and recorded whether or not they entered the inguinal canals. For voice quality, the recorded response was: 1) unbroken, 2) breaking, or 3) broken. For seminal activity, they recorded either: 1) already or 2) not yet. And for urine albumin content the recorded response was: 1) with albumin or 2) without albumin. They determined the ages at which fifty percent of the boys still displayed childhood development, and the ages at which fifty percent of the boys displayed adult development (Table 2.76).

Table 2.76

The mean age value for six sexual characteristics when 50 percent of the boys displayed childhood and 50 percent displayed adult characteristics (Hogben, Waterhouse, and Hogben, 1948)

Criteria	50% Childhood (years)	50% Adult (years)
Pubic Hair	12.9	15.0
Inguinal Canal	13.5	13.6
Voice	13.6	14.5
Axillary Hair	14.8	16.0
Mustache	15.5	unavailable
Chin Hair	16.0	unavailable

To describe a sexual maturation assessment procedure for adolescent males suitable for cross-sectional studies, University of North Carolina Professors A. Hughes Bryan and B. G. Greenberg asked a school physician to use Ellis' pubic hair scale and axillary hair scale and Greulich's genitalia scale (Bryan and Greenberg, 1952). That physician easily followed Ellis' pubic and axillary hair scales; however, he had difficulty distinguishing between the stages in Greulich's genitalia scale. Therefore, after a year's experience, the physician adopted his own three-point genitalia classifications. Based on arbitrary pubescent norms of 5.5 centimeters for penis length and 1.5 centimeters for penis width, the physician recorded the boys as: 1) immature, 2) pubescent, or 3) adolescent. To compare the boys' penis sizes with the arbitrary pubescent penis size, the physician used a penis model.

In their study of the derivation and interrelationship of adolescent males' physiological maturity indices, University of California Institute of Child Welfare Professors Arline C. Nicolson and Charles Hanley streamlined Greulich's five sexual maturation groups (Table 2.77).

Table 2.77

Five sexual maturation groups streamlined from Greulich's Maturity Groups (Nicolson and Hanley, 1953)

Stage	Description
1	Childhood penis, testes, and scrotum development
2	Penis and testes enlarge, and, lightly pigmented, downy pubic hairs appear
3	Penis lengthens, and, straight, coarse, pigmented pubic hairs intersperce with stage two pubic hair
4	Penis thickens, testes enlarges, and curly, pigmented, adult-like, pubic hairs cover most of the pubic triangle
5	Adult penis and testes, and, pubic hairs fill the pubic triangle and more

Beginning when the boys were 21 months old, the University of California Guidance Study staff annually measured 92 Berkeley, California, boys for six years, then semi-annually measured them for another ten years. Nicolson and Hanley provided the percent of their cases that reached each stage at annual age increments (Table 2.78), and from Table 2.78 the author constructed Figure 2.3.

Because he found that detailed descriptions of human bodily changes during adolescence were either widely scattered through medical journals or not described from a biological viewpoint, St. Thomas' Hospital Medical School Senior Physiology Lecturer James M. Tanner from

Table 2.78

The proportions of the boys at each age in the five sexual maturation stages (Nicolson and Hanley, 1953)

Ages (Years)	Greulich-Streamlined Sexual Maturation Stages (1)	2	3	4	5
9.0	(99.0)	1.0	0.0	0.0	0.0
9.5	(98.0)	2.0	0.0	0.0	0.0
10.0	(92.0)	8.0	0.0	0.0	0.0
10.5	(86.0)	12.0	2.0	0.0	0.0
11.0	(72.0)	23.0	3.0	2.0	0.0
11.5	(70.0)	18.0	10.0	2.0	0.0
12.0	(58.0)	18.0	20.0	5.0	0.0
12.5	(59.0)	9.0	16.0	15.0	1.0
13.0	(50.0)	3.0	22.0	23.0	1.0 (2.0)
13.5	(49.0)	3.0	16.0	31.0	9.0 (11.0)
14.0	(40.0)	2.0	9.0	16.0	22.0 (33.0)
14.5	(39.0)	2.0	4.0	10.0	12.0 (45.0)
15.0	(20.0)	0.0	3.0	7.0	25.0 (70.0)
15.5	(10.0)	0.0	3.0	3.0	14.0 (84.0)
16.0	(9.0)	0.0	0.0	2.0	5.0 (89.0)
16.5	(0.0)	0.0	0.0	3.0	6.0 (97.0
17.0	(0.0)	0.0	0.0	0.0	2.0 (100.0)
17.5	(0.0)	0.0	0.0	0.0	1.0 (100.0)
18.0	(0.0)	0.0	0.0	0.0	0.0 (100.0)

the University of London wrote *Growth at Adolescence* (Tanner, 1955). In his second chapter, Tanner discussed the development of the adolescent males' reproductive system. By combining the 1951 Stolz and Stolz text, the 1951 Reynolds and Wines report, and the 1953 Nicolson and Hanley report, Tanner created his sequence and timing figures and discussions on genitalia and pubic hair development. For his index of genitalia stages, Tanner compiled five stages and provided a pictorial guide for each of those stages (Table 2.79).

Because Tanner did not explain that he observed adolescent males to create those five genitalia stages, the author assumes that he compiled his genitalia stages from the references that he cited. For his index of pubic hair, Tanner compiled five stages and provided

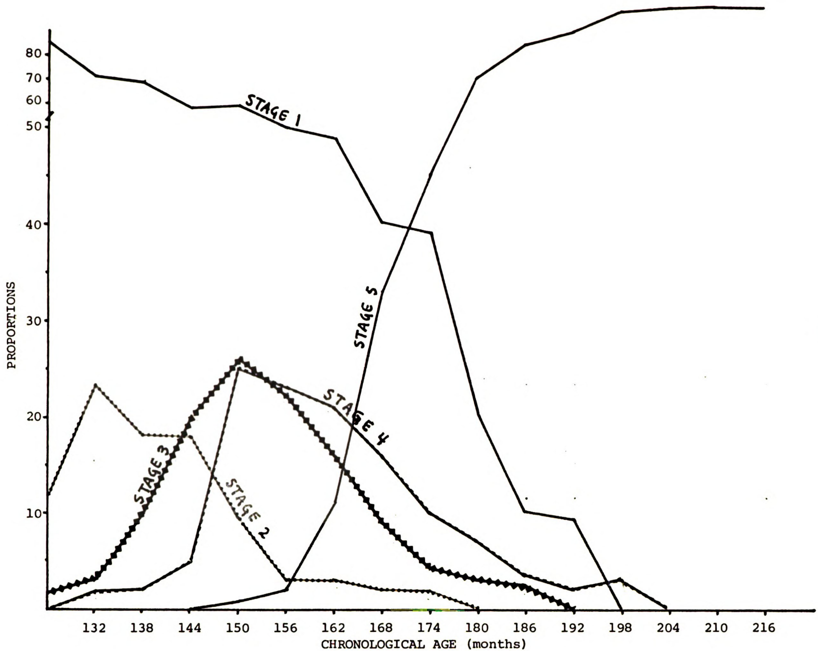


Figure 2.3 Observed sexual maturation stages proportions (Nicholson and Hanley, 1953)

Table 2.79

Genitalia development stages for adolescent males (Tanner, 1955)

Stage	Description
1	Early childhood testes, scrotum, and penis size and proportion
2	Testes enlarged. Scrotum reddens and changes texture. Little or no penis enlargement.
3	Penis lengthens. Testes and scrotum grew more.
4	Penis widened and its glans developed. Testes and scrotum further enlarged. Scrotum darkened.
5	Adult size and shape genitalia.

a pictorial guide for each of those stages (Table 2.80). Again, the author assumes that Tanner compiled those pubic hair stages from indices that others created. Tanner did not provide indices for axillary hair or facial hair.

Table 2.80

Pubic hair stages for adolescent males (Tanner, 1955)

Stage	Description
1	The pubes vellus is identical to the abdominal wall vellus, or, no pubic hair development
2	A sparse growth of long, slightly pigmented, straight or slightly curled, downy, pubic hairs appeared at the base of the penis
3	Considerably darker, coarse, curled pubic hairs spread over the pubic junction
4	Adult-type pubic hairs cover only the pubic triangle
5	Adult-type and quantity pubic hair

Using the data Phillips Academy Physician J. Roswell Gallagher helped collect for W. W. Greulich's classic 1942 report, Harvard Medical School Pediatrics Department Professor Edward E. Hunt, Boston's Forsythe Children's Dental Infirmary worker Grace Cooke, and Harvard University Anthropology Professor J. Roswell Gallagher compared various researchers' data on sexual maturation (Hunt, Cooke, and Gallagher, 1958). From Greulich's five Maturity Groups and four intermediate groupings, they classified boys into nine sexual maturation groups. To create their sexual maturation stages, Hunt, Cooke, and Gallagher streamlined Greulich's five Maturity Groups.

By statistically analyzing Reynolds' and Wines' data, Tanner's data, and their own data, Hunt et al. tested the reliability of Greulich's five Maturity Groups. They concluded that on a cross-sectional basis, observers judged pubic hair initiation more accurately than they judged the advent of testes growth. They hypothesized that the difficulty that the observers had in recognizing the advent of testes growth resulted from the absence of comparative photographs when the boys were pre-pubescent. When the observers assessed Greulich's Maturity Group three, they found satisfactory agreement. Because 1) Greulich's fifth stage does not describe sexual maturation and 2) the phenomenon of penis shrinkage in late adolescence confused some observers, they concluded that on a cross-sectional basis Greulich's Maturity Group five was very difficult to assess.

To study the future growth of short-statured 14 year old males, Harvard Medical School Pediatrics Professors J. Roswell Gallagher and Carl C. Seltzer X-rayed the hand-wrist area and used Greulich's Maturity Groups to assess the sexual maturation of 101 boys aged 14 years old annually for three years (Gallagher and Seltzer, 1961). They used the Bayley-Pinneau skeletal maturation tables for predicting the boys' adult height, and compared their sexual maturation to their predicted adult height.

They determined that 14 year old boys less than 63 inches tall differed in rate of adolescent growth from 14 year old boys more than 63 inches tall. For the nineteen 14 year old boys less than 63 inches tall, they found an average height gain of six and one-half inches, whereas for the eighty-two 14 year old boys more than 63 inches tall, they found an average height gain of only four inches. Of those nineteen 14 year old boys less than 63 inches tall, the seven rated at Maturity Group two averaged 7.1 inches in actual height gain, with a range of 6.0 to 8.3 inches, the seven rated at Maturity Group three averaged 6.8 inches in actual height gain, with a range of 5.7 to 8.3 inches, the four rated at Maturity Group four averaged 5.5 inches in actual height gain, with a range of 5.1 to 5.9 inches, and the one rated at Maturity Group five grew only 3.2 inches more. When physicians 1) cannot obtain a hand-wrist X-ray, 2) do not know how to read the hand-wrist X-ray, or 3) do not have the time to read the X-ray and work with the Bayley-Pinneau tables for predicting adult height, they advised that Greulich's Maturity Groups may prove helpful in predicting the future height gain potential of adolescent males.

As one of the several measurements that he advised physicians to take on their adolescent male patients, Physician Frank Falkner included a test for pubertal status (Falkner, 1961). In the reference section of his paper, Falkner cited Tanner's *Growth at Adolescence*. However, within the paper, Falkner did not credit anyone for the creation of his genital and pubic hair development stages (Tables 2.81 and 2.82). By combining the boys' genital and pubic hair development ratings, Falkner created his "Pubertal Status" score. Therefore, if

Table 2.81

Genital development stages for adolescent males (Falkner, 1961)

Stage	Description
1	Childhood penis and testes size
2	Testes enlarge and scrotal skin reddens and coarsens
3	Penis lengthens
4	Penis general size enlargens and scrotal skin darkens
5	Adult genitalia

Table 2.82

Pubic hair development stages for adolescent males (Falkner, 1961)

Stage	Description
1	No true pubic hair
2	Sparse, downy pubic hair at the base of the penis
3	Pigmented, coarse, curly, pubic hair increases in amount
4	Adult pubic hair limited to the pubic triangle
5	Adult pubic hair spread to the thighs
6	Adult pubic hair spread up to the linea alba

he rated a boy's genital stage as 3 and his pubic hair stage as 2, then he recorded the boy's pubertal status as 3:2.

To determine whether or not adolescent boys with advanced sexual maturation had an academic advantage over adolescent boys with delayed sexual maturation, J. W. B. Douglas and J. M. Ross reviewed 1,872 school medical examinations of 15 year old boys wherein the school physician recorded the boys' genitalia maturation, pubic hair development, axillary hair development, and vocal quality (Douglas and Ross, 1964). Based upon those primary and secondary sexual characteristics, Douglas and Ross created four classifications of puberty (Table 2.83). They reported the number of the 1,872 fifteen year old males that they rated in each of their four classifications of

Table 2.83

Four classifications of puberty for adolescent males (Douglas and Ross, 1964)

Classification	Description
Infantile	Infantile genitalia, no pubic hair, no axillary hair, and an unbroken voice
Early Signs	Early adolescent genitalia, some pubic hair, and some axillary hair
Advanced	More maturation than the Early Signs category, but not adult maturation
Mature	Genitalia completely developed; penis and glans substantially enlarged, testes enlarged, scrotum pendulous and rugose, pubic hair profuse, axillary hair profuse, and, a broken voice

puberty (Table 2.84). They did not explain why they placed 177 boys in an unknown puberty classification.

Table 2.84

The proportion of the boys in the four classifications of puberty (Douglas and Ross, 1964)

Classification	Actual Number	Proportion
Infantile	177	9.5
Early Signs	597	31.9
Advanced	504	26.9
Mature	417	22.3
Unknown	177	9.5

To determine the sexual maturation process of Hong Kong Chinese adolescent boys, University of Hong Kong Anatomist K. S. Frances Chang, Physician Marjorie M. M. C. Lee, and S. J. Chan assessed the penis growth, pubic hair development, axillary hair development, and vocal quality of 3,658 Southern Chinese boys aged 6 to 18 years old (Chang, Lee, and Chan, 1966). They created five genital development stages (Table 2.85) and five pubic hair development stages

Table 2.85

Five genitalia development stages for adolescent males
(Chang, Lee, and Chan, 1966)

Stages	Descriptions
1	Childhood size penis
2	Penis a little larger than childhood size
3	Penis has lengthened and widened
4	Penis has enlarged further and its skin has darkened
5	Adult size and shape penis

(Table 2.86). They determined the age of the boys at the point where fifty percent of the boys achieved each genitalia and pubic hair stage. As a result of their discussion, the author created Table 2.87, which shows the ages at which 50 percent of the boys attained

Table 2.86

Five pubic hair stages for adolescent males
(Chang , Lee, and Chan, 1966)

Stages	Description
1	No pubic hair present
2	Sparse, straight or slightly curled, downy pubic hair present only at the base of the penis
3	Numerous, dark, coarse, and curly pubic hairs spread slightly beyond the base of the penis
4	Adult-type pubic hairs present, but, limited in distribution
5	Adult-type pubic hairs spread to the thighs, up the linea alba, and down to the scrotum

Table 2.87

The ages at which 50 percent of the boys attained the genitalia and pubic hair stages (Chang , Lee, and Chan, 1966)

Genital Stages	50% Age	Pubic Hair Stages	50% Age
2	13.19	2	13.31
3	13.75	3	13.70
4	14.60	4	14.60
5	16.47	5	16.47

the genitalia and pubic hair stages. They concluded that the ages of onset and duration of genital and pubic hair stages for their subjects were about the same. By calling stage 1 non-pubescent, by combining stages 2, 3, and 4 and calling them pubescent, and by calling stage 5 post-pubescent, they constructed graphs to show the relationship between those three sexual maturation phases in adolescent males. From the graphs of Chang, Lee, and Chan, the author estimated the proportion of boys in each phase from ages 11 through 18 for genitalia and pubic hair (Tables 2.88 and 2.89). Figures 2.4 and 2.5 graphically illustrate Tables 2.88 and 2.89.

Table 2.88

The proportion of the boys at each age in the three genitalia phases (Chang, Lee, and Chan, 1966)

Age	Non-Pubescent	Pubescent	Post-Pubescent
11.0	99.0	1.0	0.0
12.0	88.0	12.0	0.0
13.0	57.0	42.0	1.0
14.0	13.0	82.0	5.0
15.0	3.0	85.0	13.0
16.0	1.0	63.0	36.0
17.0	-0.0	33.0	67.0
18.0	0.0	13.0	87.0

Table 2.89

The proportion of boys at each age in the three pubic hair phases (Chang, Lee, and Chan, 1966)

Age	Non-pubescent	Pubescent	Post-pubescent
11.0	99.0	1.0	0.0
12.0	90.0	10.0	0.0
13.0	60.0	39.0	1.0
14.0	17.0	80.0	3.0
15.0	4.0	83.0	13.0
16.0	1.0	60.0	39.0
17.0	0.0	32.0	68.0
18.0	0.0	16.0	84.0

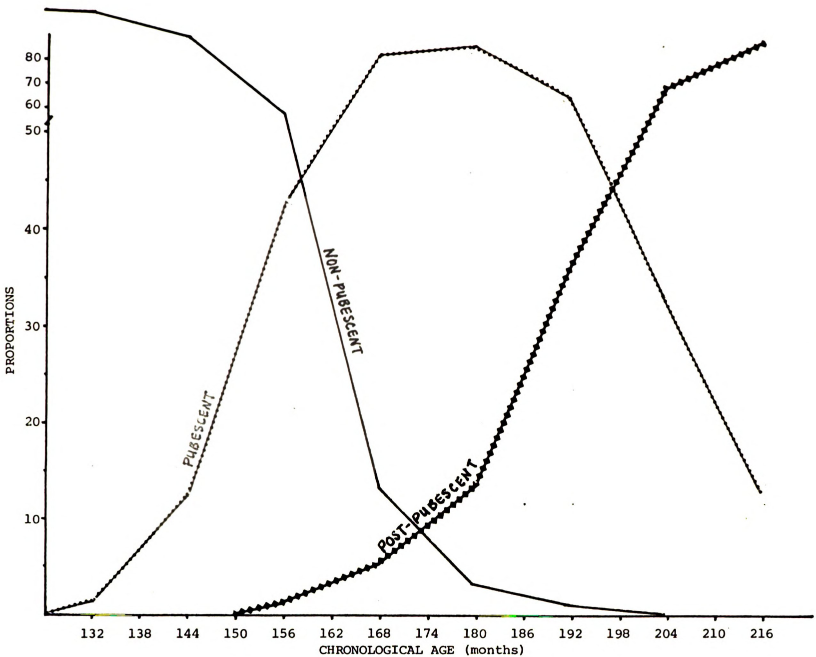


Figure 2.4 Observed genitalia phase proportions (Chang, Lee, and Chan, 1966)

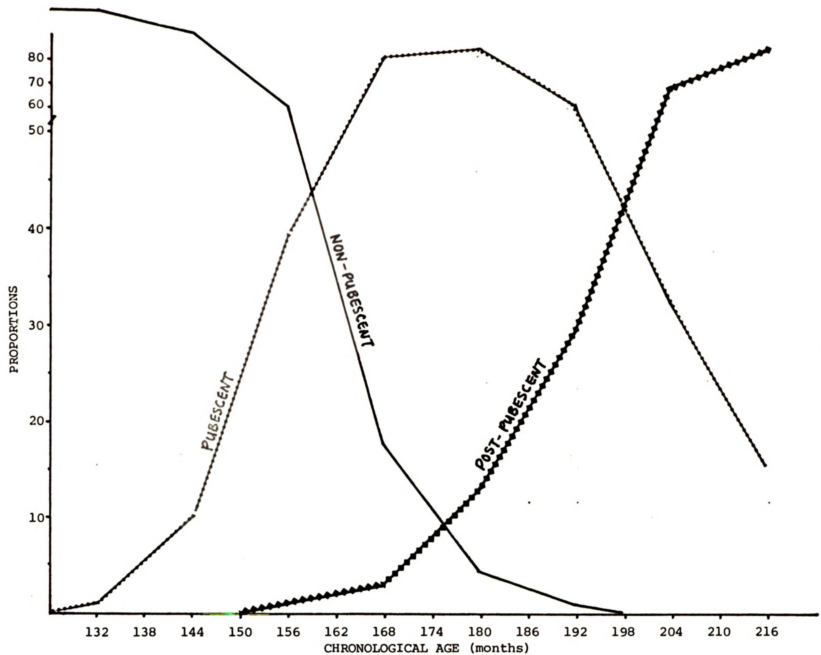


Figure 2.5 Observed pubic hair proportions (Chang, Lee, and Chan, 1966)

Chang et al. found that axillary hair appeared when genitalia and pubic hair reached their fourth stages. They found it difficult to determine vocal changes from cross-sectional sampling, but when they did assess vocal change they determined that those changes occurred when genitalia and pubic hair reached their third stage.

Relying upon the pioneer works of Greulich et al., Ellis, and Reynolds and Wines to classify the sexual maturation of adolescent boys, Michigan State University Masters degree candidate Michael G. Marshall created five checklists for the primary and secondary sexual characteristics of facial hair development, axillary hair development, pubic hair development, penis development, and scrotum development (Tables 2.90 through 2.94) (Marshall, 1967).

Table 2.90

Facial hair development checklist (Marshall, 1967)

Stage	Description (Ask the boys if they shaved today)
1	No facial hairs on mustache, sideburns, or, chin
2	Unpigmented sideburn hair present
3	Downy, lightly pigmented hairs present at the corners of the upper lip
4	Lightly pigmented, coarse, mustache hairs present
5	Adult mustache hairs present

Table 2.91

Axillary hair development checklist (Marshall, 1967)

Stage	Description
1	No axillary hairs present
2	Slight, unpigmented axillary hairs present
3	Downy, lightly pigmented axillayr hairs in one or both axilla
4	Pigmented, coarse, axillary hairs cover ½ to 1 inch area
5	Coarse, curly, well-developed axillary hairs fill both axilla

Table 2.92

Pubic hair development checklist (Marshall, 1967)

Stage	Description
1	No pubic hairs present
2	Downy, unpigmented, pubic hairs present
3	Sparse, pigmented, straight, pubic hairs present
4	Curly, coarse, pigmented, pubic hairs present, but not extensive
5	Adult-type pubic hairs extends to the medial surface of the thighs

Table 2.93

Penis development checklist (Marshall, 1967)

Stage	Description
1	Childhood sized penis
2	Slight penis growth
3	Penis shaft lengthens, with the glans proportionately small
4	Penis shaft lengthens and thickens
5	Maximum size penis

Table 2.94

Scrotum and testes development checklist (Marshall, 1967)

Stage	Description
1	Childhood scrotum and testes
2	Scrotum and testes visibly enlarged, one testicle may show lowered
3	One testicle definitely has lowered
4	Scrotum darkening beginning
5	Maximum size testes and maximum scrotum darkening

Because he wanted Junior High School Boys Physical Education teachers to assess the students' sexual maturation on a cross-sectional basis for each characteristic, Marshall described each stage independent of the other stages. After inspecting the data obtained from the 671 East Lansing Junior High School boys that the physical education teachers had assessed, Marshall combined the five sexual characteristics scores into a single Sexual Maturation Indicator Value (SMIV). To calculate the SMIV score, he used the boys' pubic hair development score as the primary indicator and used the facial hair, axillary hair, penis and scrotum and testes scores to add to or subtract from the pubic hair rating (Table 2.95). Once he obtained

Table 2.95

SMIV formula (Marshall, 1967)

SMIV = (Facial hair rating + Axillary hair rating + Penis development rating + Scrotum and testes development rating) - (4 X Pubic hair rating)/ 10 + Pubic hair rating

the boys' SMIV score, he subdivided the SMIV scores into five puberty groups (Table 2.96). Marshall reported the number of subjects, their average age, and the standard deviation for each of his puberty groups (Table 2.97).

Table 2.96

Five puberty groups by the SMIV breakdown (Marshall, 1967)

Puberty Group	SMIV Breakdown
A	1.0 to 1.5
B	1.6 to 2.5
C	2.6 to 3.5
D	3.6 to 4.5
E	4.5 to 5.0

To determine the extent to which the pubertal ratings of adolescent males were associated with each primary and secondary sexual characteristics and to determine their between and within rater reliabilities, Harvard School of Public Health Professor H. Bontourline

Table 2.97

The number, mean values and standard deviations of the boys in the five puberty groups (Marshall, 1967)

Descriptive Data	Marshall's Puberty Groups				
	A	B	C	D	E
Number	118	209	206	132	11
Mean(months)	154.17	156.61	161.20	165.21	172.09
Standard Deviation	7.16	6.66	6.90	7.71	6.95

Young, Stanford University Anatomist W. W. Greulich, Harvard Medical School Professors J. R. Gallagher and T. Cone, and Washington, D.C., Children's Hospital Physician F. Heald assessed eleven primary and secondary sexual characteristics of 70 Boston area adolescent males three separate times at two month intervals (Young, Greulich, Gallagher, and Heald, 1968). Young selected those 70 boys to represent a rectangular distribution of sexual development rather than a normal bell-shaped distribution. Using the descriptions developed by Greulich (1942), Schonfeld (1943), and Stolz and Stolz (1951), they created six classifications of pubertal maturity (Table 2.98).

To assess the eleven primary and secondary sexual characteristics necessary to classify the boys in their six pubertal maturity stages, Young et al. used Tanner's genitalia pictorial guide and several checklists. Tables 2.99, 2.100 and 2.101 contain their pubic hair, axillary hair, and facial hair checklists. On their checklists, they rated the boys at plus, minus, or right at each stage. Thereby, they had sixteen ratings for each checklist.

In addition to their pubic hair checklist, they categorized the pubic hair configuration as: 1) concave, 2) straight, and 3) convex. In addition to their facial hair checklist, they asked the boys if they

Table 2.98

Six classifications of pubertal maturity for adolescent males (Young, Greulich, Gallagher, and Heald, 1968)

Classification	Description
1 (Childhood)	Testicular volume 1 No secondary sexual characteristics present
2 (Prepuberal)	Testicular volume 2 Pubic hair 1 Axillary and body hair 0 Slight increase in penis length and diameter Eyebrows, line of cheek, hairline, and voice 1
3 (Puberty 1st. Stage)	Testicular volume 3 or 4 Pubic hair 2 Axillary hair 0 Penis definitely enlarged Facial and body hair 1 and 2
4 (Puberty 2nd Stage)	Testicular volume 4 Pubic hair 3 or 4 Axillary hair 2 or 3 Penis further enlarged
5 (Puberty 3rd Stage)	Testicular volume 5 Pubic hair 3 or 4 Axillary hair 2 or 3 Penis further enlarged
6 (Adult form)	Testicular volume 6 Adult voice Pubic hair 4 Axillary hair 3 or 4

Table 2.99

Pubic hair checklist (Young, Greulich, Gallagher, and Heald, 1968)

Stage	Description
0	No pubic hairs present
1	Downy, unpigmented, fine, straight, pubic hairs present
2	Pigmented, coarse, crinkled, pubic hair cover a small area
3	Pigmented, coarse, crinkled, pubic hairs cover a moderate area
4	Pigmented, coarse, crinkled, pubic hiars cover a considerable area

sh

ti

an

Be

as

tes

co

re

the

(Ta

Table 2.100

Axillary hair checklist (Young, Greulich, Gallagher, and Heald, 1968)

Stage	Description
0	No axillary hairs present
1	Fine, straight, axillary hairs cover a small area
2	Pigmented, coarse, curly or crinkled, axillary hairs cover a small area
3	Pigmented, coarse, curly or crinkled, axillary hairs cover a moderate area
4	Pigmented, coarse, curly or crinkled, axillary hair cover a considerable area

Table 2.101

Facial hair checklist (Young, Greulich, Gallagher, and Heald, 1968)

Stage	Description
0	No facial hairs present
1	Downy, straight, facial hairs present
2	Some pigmented, coarse, facial hairs present
3	Pigmented, coarse, facial hairs present in moderate amounts
4	Pigmented, coarse, facial hairs present in considerable amounts

shaved. Upon receiving an affirmative response, they asked how many times per week the boys shaved. For their assessment of penile length and diameter, they measured the penis with a millimeter metal tape. Because they referred to the 1943 report by Schonfeld, the author assumes that they measured the stretched but flaccid penis. For their testicular volume measurement, they adopted Schonfeld's technique of comparing the testicle with the known volume of wooden models. They reported testicular volumes, penis length, and penis diameters that they considered necessary to rate their boys in stages 1 through 6 (Table 2.102). Additionally, they provided checklists for rating the

Table 2.102

The testicular volumes, penis lengths, and penis diameters necessary for the six puberty ratings (Young, Greulich, Gallagher, and Heald, 1968)

Pubertal Rating	Testicular Volume (ccm)	Penis Length (cm)	Penis Diameter (cm)
1	1.5	1.6	1.3
2	3.0	2.1	1.6
3	6.5	2.8	2.1
4	10.5	3.2	2.5
5	15.5	3.5	2.9
6	21.5	4.0	3.2

breast development (Table 2.103) and the eyebrow development (Table 2.104).

Table 2.103

Breast development checklist for adolescent males (Young, Greulich, Gallagher, and Heald, 1968)

Stage	Description
0	No visible or palpable mammary gland
1	Slight, but definite palpable mammary gland
2	More marked mammary gland enlargement
3	Moderate to considerable mammary gland enlargement
4	Considerable bilateral mammary gland enlargement

Table 2.104

Eyebrow development checklist for adolescent males (Young, Greulich, Gallagher, and Heald, 1968)

Stage	Description
1	Infantile eyebrows present
2	Eyebrows thicken
3	Eyebrow hairs coarsen, and, the eyebrows thicken and grow towards the midline
4	Adult-type eyebrows present

By comparing the scalp hairline to four hairline diagrams that ranged from bowlike to double apical, they rated the hairline on a four-point scale. Similarly, by comparing the facial configuration to four jawline diagrams that ranged from round-chin to V-shaped chin jawlines, they rated the facial configuration on a four-point scale. For their measurement of vocal quality, they categorized the voices as: 1) childish, 2) broken, 3) adult.

They reported the test-retest correlations for the five raters on the eleven primary and secondary sexual characteristics assessments (Table 2.105). The average test-retest correlations for facial hair,

Table 2.105

The test-retest correlations for the five observers in this study (Young, Greulich, Gallagher, and Heald, 1968)

Sexual Characteristics	Observers					
	1	2	3	4	5	AVE
Eyebrows	.80	.80	.82	.83	.81	.812
Voice	.86	.85	.86	.87	.86	.870
Facial hair	.86	.87	.88	.87	.85	.864
Axillary hair	.91	.90	.91	.91	.89	.904
Body hair	.79	.81	.80	.80	.77	.794
Pubic hair	.96	.95	.96	.96	.96	.958
Hair line	.77	.77	.78	.78	.77	.774
Cheek line	.85	.86	.86	.87	.87	.862
Pubic hair shape	.65	.64	.64	.64	.61	.636
Penis length	.83	.81	.82	.84	.84	.828
Testicular volume	.95	.94	.93	.97	.96	.954

axillary hair, pubic hair, and testicular volume of .864, .904, and .954, respectively, suggest that specialists in child growth and development should be able to assess those sexual characteristics with acceptable reliability.

From those test-retest correlations, Young et al. concluded that testicular volume, pubic hair, and axillary hair demonstrated sufficient assessment reliability for them to use as major variables in assessing sexual development. To determine the beta weight that each observer attributed to those three major variables, they computed separate multiple regression analyses of each observer's assessments. They reported each observer's standardized partial regression coefficients for their selected three major variables (Table 2.106).

Table 2.106

The standardized partial regression coefficients for five observers (Young, Greulich, Gallagher, and Heald, 1968)

Sexual Characteristic	Observers					
	1	2	3	4	5	AVE
Axillary hair	.282	.283	.309	.264	.180	.267
Pubic hair	.449	.478	.467	.387	.403	.434
Testicular volume	.295	.252	.245	.378	.434	.331

By multiplying the averaged standardized partial regression coefficients by the standard deviation of the pubertal maturity ratings and then dividing the value by the standard deviation of each major variable, they determined a regression equation for pubertal age. The resulting regression equation for predicting the Pubertal Age of Adolescent Males was: $(P\hat{A}AM)$ = (.5 + .5[pubic hair] + 4[testicular volume] + .3[axillary hair]).

At this point they stated that to evaluate the pubertal age regression equation, they needed a criterion separate from sexual development that correlated highly with biological maturity. They chose the Greulich-Pyle Atlas to assess skeletal age.

From 120 adolescent males from Palermo, Italy, Young selected 42 boys whose skeletal ages ranged from 11 to 16 years on whom they had the data from which they could determine pubertal age. They reported the correlations between the boys' chronological age, skeletal age, first pubertal maturity rating, calculated puberty age, and second pubertal maturity rating (Table 2.107). The author

Table 2.107

The correlations between chronological age, skeletal age, a first pubertal maturity rating, a calculated pubertal age, and a second pubertal maturity rating (Young, Greulich, Gallagher, and Heald, 1968)

	Chrono. Age	Skeletal Age	1st Pubertal Maturity Rat.	Cal. Pub. Age	2nd. P.M.T.
Chronological Age	1.00				
Skeletal Age	.53	1.00			
1st Pubertal Mat. Rating	.54	.92	1.00		
Calc. Pubertal Age	.52	.92	.99	1.00	
2nd Pubertal Mat. Rating	.53	.93	.99	.99	1.00

noted that the correlations between chronological age and their estimates of pubertal age were low (.54 and less), but because they derived their calculated pubertal age from their pubertal maturity ratings, those correlations were high (.99). Although their specially-selected rectangular sample partially explains their high correlation (.92) between skeletal age and pubertal age, they concluded that their equation for pubertal age provided a precise method for calculating the pubertal maturity of adolescent boys.

To determine the chronological age variation at which adolescent boys start and pass through Tanner's genitalia and pubic hair stages, University of London's Institute of Child Health, Growth, and Development Professors W. A. Marshall and J. M. Tanner viewed the nude, full body photographs of 228 British boys who were subjects in the Harpenden

Growth Study (Marshall and Tanner, 1970). The Institute collected those data every three months during the adolescent age period. By sequentially viewing each boy's genitalia and pubic hair changes, and by using Tanner's genitalia and pubic hair stages, they classified the maturational stages of the boys. They considered their pubic hair stage 2 as an unreliable rating stage because unpigmented hairs did not show well on the photographs. If the pubic hair had spread up the linea alba, they assessed those boys as stage 6. They calculated the average age and standard deviation at which the boys reached each stage (Table 2.108). They reported the proportion of

Table 2.108

The average age and standard deviation that the boys attained the genitalia and pubic hair stages (Marshall and Tanner, 1970)

Descriptive Data	Genitalia Stages				Pubic Hair Stages			
	2	3	4	5	2*	3	4	5
Mean	11.64	12.85	13.77	14.92	13.44	13 90	14.36	15.18
S.D.	1.07	1.04	1.02	1.10	1.09	1.04	1.08	1.07

(*They considered this stage unreliable.)

boys in the five genitalia stages while also in each of the pubic hair stages (Table 2.109), and conversely, the proportion of boys in the five pubic hair stages while also in each of the genitalia stages (Table 2.110). They noted that forty-one percent of their boys reached genitalia stage 4 before their pubic hair development began. If the boys' genitalia development paralleled their pubic hair development, then the proportion of the boys would be greatest when the genitalia and pubic hair stages were equal. However, the cell containing the genitalia and pubic hair stage 2 boys provided only thirteen percent

Table 2.109

The proportion of the boys in the five genitalia stages across the four pubic hair stages (Marshall and Tanner, 1970)

Pubic Hair	Tanner's Genitalia Stages			
Stages	2	3	4	5
1	98.0	72.0	16.0	0.0
2	2.0	17.0	37.0	0.0
3	0.0	9.0	36.0	8.0
4	0.0	2.0	11.0	54.0
5	0.0	0.0	0.0	38.0

Table 2.110

The proportion of the boys in the five pubic hair stages across the four genitalia stages (Marshall and Tanner, 1970)

Genitalia	Tanner's Pubic Hair Stages			
Stages	2	3	4	5
1	1.0	0.0	0.0	0.0
2	13.0	4.0	0.0	0.0
3	45.0	17.0	6.0	0.0
4	41.0	75.0	65.0	10.0
5	0.0	4.0	29.0	90.0

of the sample. Likewise, the cell containing the genitalia and pubic hair stage 3 boys provided only seventeen percent of the sample. However, the genitalia and pubic hair stage 4 cell provided sixty-five percent of the sample, and the cell for the genitalia and pubic hair stage 5 boys provided ninety percent of the boys. Relative to rate of genitalia development after reaching stage 3, the boys' rate of pubic hair development increased.

To establish the ages at which Istanbul adolescent males attained their secondary sexual development stages, Istanbul University Pediatrics and Administrative Services Professors Olcay Neyzi, Hulpa Alp, Ayfer Yalcindag, Suleyman Yakacikli, and Alper Orphon studied

1,5

an

bi

ha

ti

Af

te

ow

de

ad

at

ti

1,530 Turkish boys aged 9 to 17 years (Neyzi, Alp, Yalcindag, Yakacikli, and Orphon, 1975). All of the boys were within three months of their birthday. By visually inspecting the boys' penis, facial hair, axillary hair, and pubic hair and by palpating the boys' testes and breast tissue, Suleyman Yakacikli assessed the boys' sexual development. After reviewing the Reynolds and Wines 1951 report, the Tanner 1962 text, and the Marshall and Tanner 1970 report, they adopted their own rating scales for assessing pubic hair development, facial hair development, axillary hair development, and laryngeal development of adolescent males (Tables 2.111 through 2.114). They reported the mean attainment age that their boys reached their four sexual characteristics' development stages (Table 2.115).

Table 2.111

Pubic hair stages (Neyzi et al., 1975)

Stage	Description
PH1	No pubic hair present
PH2	Sparse, straight or slightly curly, pubic hairs present at the base of the penis
PH3	Darker, coarser, curlier, pubic hairs spread over the pubis junction
PH4	Adult-type pubic hairs cover only the pubis, not the inner thigh area
PH5	Adult-type pubic hairs cover adult or near-adult area

Table 2.112

Facial hair stages (Neyzi et al., 1975)

Stage	Description
Ax1	No axillary hairs present
Ax2	Sparse axillary hairs present
Ax3	Moderate amount of curly axillary hairs present
Ax4	Abundant amount of curly axillary hairs present

Table 2.113

Facial hair stages (Neyzi et al., 1975)

Stages	Description
FH1	No facial hairs present
FH2	Slightly long, pigmented facial hairs present at the corners of the upper lip
FH3	Complete mustache present and some facial hairs on the cheeks and chin
FH4	Adult or near-adult facial hairs distribution and quantity present.

Table 2.114

Laryngeal stages (Neyzi et al., 1975)

Stages	Description
V1	Infantile vocal quality present
V2	Some voice cracking present
V3	Noticeable male quality voice present
V4	Definite male quality voice present

Table 2.115

The mean attainment age that the boys reached the pubic hair, axillary hair, and laryngeal stages (Neyzi et al., 1975)

	Pubic Hair Stages			Axillary Hair Stages		Facial Hair Stages		Larynx Stages	
Means	PH2	PH3	PH4	AX2	AX3	FH2	FH3	V2	V3
Ages	12.28	13.81	16.07	13.55	15.78	14.72	X	13.71	16.22

In their appendix, Neyzi et al. reported the proportion of the boys at each age that attained each stage of the four sexual characteristics (Tables 2.116 through 2.119). In parentheses, the author placed the proportion value. From those tables, the author constructed Figures 2.6 through 2.9 (Neyzi, Alp, Yalcindag, Yakacikli, and Orphon, 1975).

Table 2.116

The proportion of the boys at each age in the five pubic hair stages (Neyzi et al., 1975)

AGE (Years)	N	Pubic Hair Stages				
		PH1	PH2	PH3	PH4	PH5
9	167	160(95.8)	7(4.2)	0(0.0)	0(0.0)	0(0.0)
10	175	160(91.4)	14(8.0)	1(0.6)	0(0.0)	0(0.0)
11	195	162(83.1)	30(15.4)	3(1.5)	0(0.0)	0(0.0)
12	163	100(61.3)	51(31.2)	11(6.7)	1(0.6)	0(0.0)
13	164	60(36.6)	62(37.8)	34(20.7)	8(4.9)	0(0.0)
14	162	35(21.6)	46(28.4)	63(38.9)	17(10.5)	1(0.6)
15	162	7(4.3)	24(14.8)	81(50.0)	38(23.5)	12(7.4)
16	177	0(0.0)	14(7.9)	64(36.2)	78(44.1)	21(11.9)
17	165	0(0.0)	4(2.4)	54(32.7)	70(42.4)	37(22.4)

Table 2.117

The proportion of the boys at each age in the four axillary hair stages (Neyzi et al., 1975)

AGE (Years)	N	Axillary Hair Stages			
		Ax1	Ax2	Ax3	Ax4
9	167	167(100.0)	0(0.0)	0(o.0)	0(0.0)
10	175	174(99.4)	1(0.6)	0(0.0)	0(0.0)
11	195	191(97.9)	3(1.5)	1(0.5)	0(0.0)
12	163	144(88.3)	16(9.8)	3(1.8)	0(0.0)
13	164	117(71.3)	40(23.4)	7(4.3)	0(0.0)
14	162	66(40.7)	77(47.5)	18(11.1)	1(0.6)
15	162	20(12.3)	89(54.9)	42(25.9)	11(6.8)
16	177	2(1.1)	74(41.1)	78(44.1)	23(13.0)
17	165	0(0.0)	48(29.1)	79(47.9)	38(23.0)

Table 2.118

The proportion of the boys at each age in the four facial hair stages (Neyzi et al., 1975)

AGE (Years)	N	Facial Hair Stages FH1	FH2	FH3	FH4
9	167	167(100.0)	0(0.0)	0(0.0)	0(0.0)
10	175	175(100.0)	0(0.0)	0(0.0)	0(0.0)
11	192	191(99.5)	1(0.5)	0(0.0)	0(0.0)
12	163	160(98.2)	3(1.8)	0(0.0)	0(0.0)
13	164	143(87.2)	20(12.3)	1(0.6)	0(0.0)
14	162	108(66.7)	51(31.5)	2(1.2)	1(0.6)
15	162	67(41.4)	76(46.9)	19(11.7)	0(0.0)
16	177	41(23.2)	88(49.7)	46(26.0)	2(1.1)
17	165	13(7.9)	62(37.6)	83(50.3)	7(4.2)

Table 2.119

The proportion of the boys at each age in the four laryngeal stages (Neyzi et al., 1975)

AGE (Years)	N	Laryngeal Stages V1	V2	V3	V4
9	167	167(100.0)	0(0.0)	0(0.0)	0(0.0)
10	175	173(98.9)	2(1.1)	0(0.0)	0(0.0)
11	195	191(97.9)	4(2.1)	0(0.0)	0(0.0)
12	163	147(90.2)	15(9.2)	1(0.6)	0(0.0)
13	164	120(73.2)	40(24.4)	4(2.4)	0(0.0)
14	162	80(49.4)	75(46.5)	7(4.3)	0(0.0)
15	162	34(21.0)	90(55.6)	36(22.2)	2(1.2)
16	177	10(5.6)	83(46.9)	82(46.3)	2(1.1)
17	165	2(1.2)	44(26.7)	102(61.8)	17(10.3)

Summary

From the preceding review, the author has drawn the following principles of sexual development that are pertinent to this investigation: 1) when researchers cannot obtain hand-wrist X-rays from which they can determine the biological age of adolescent males, the next best predictor of their biological age is an assessment of the boys' sexual development, 2) the adolescent males' primary sexual

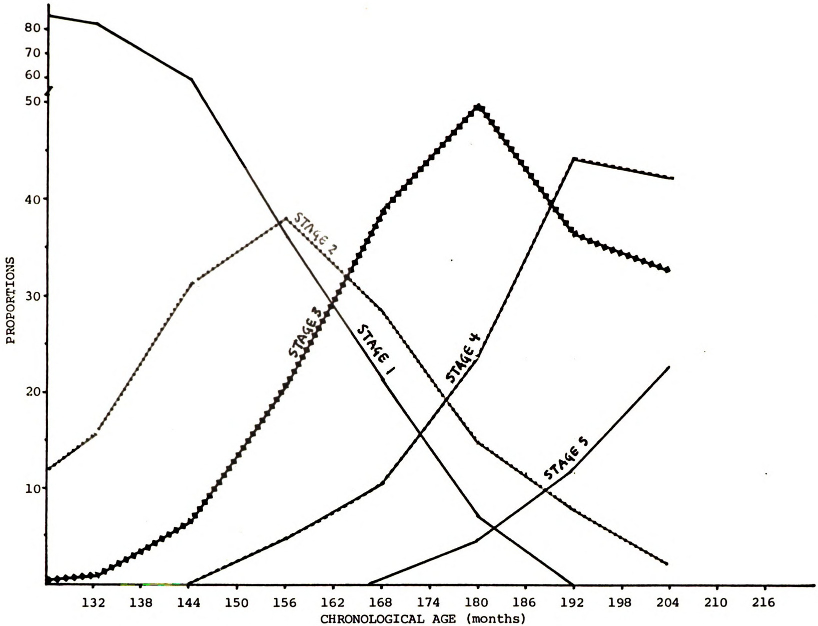


Figure 2.6 Observed pubic hair stage proportions (Neyzi et al, 1975)

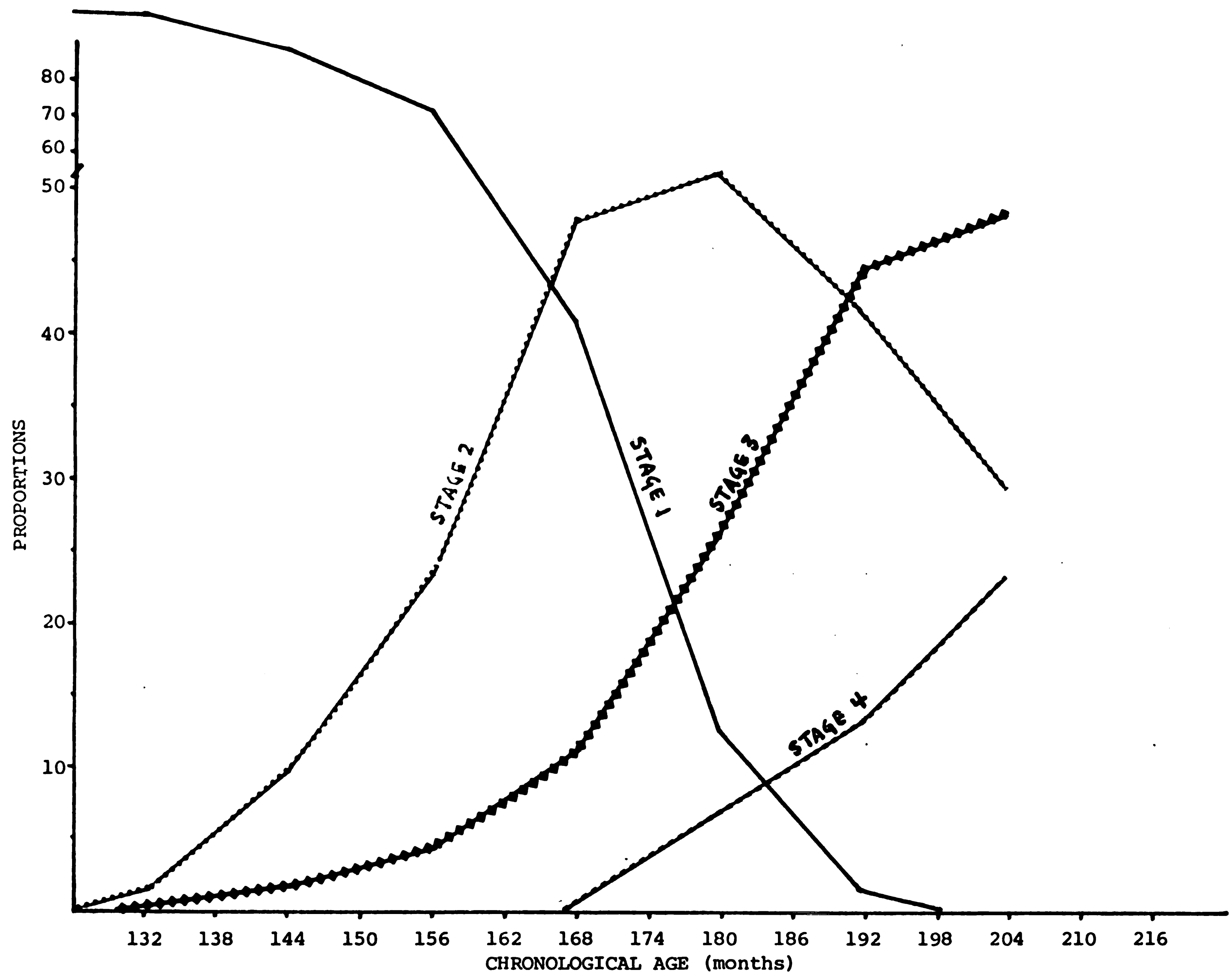


Figure 2.7 Observed axillary hair stage proportions (Neyzi et al, 1975)

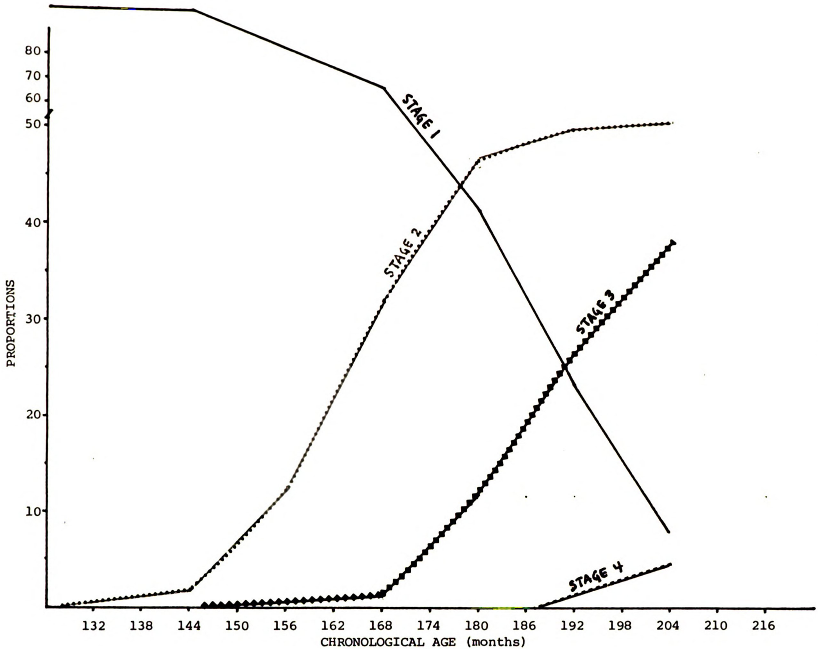


Figure 2.8 Observed facial hair stage proportions (Neyzi et al, 1975)

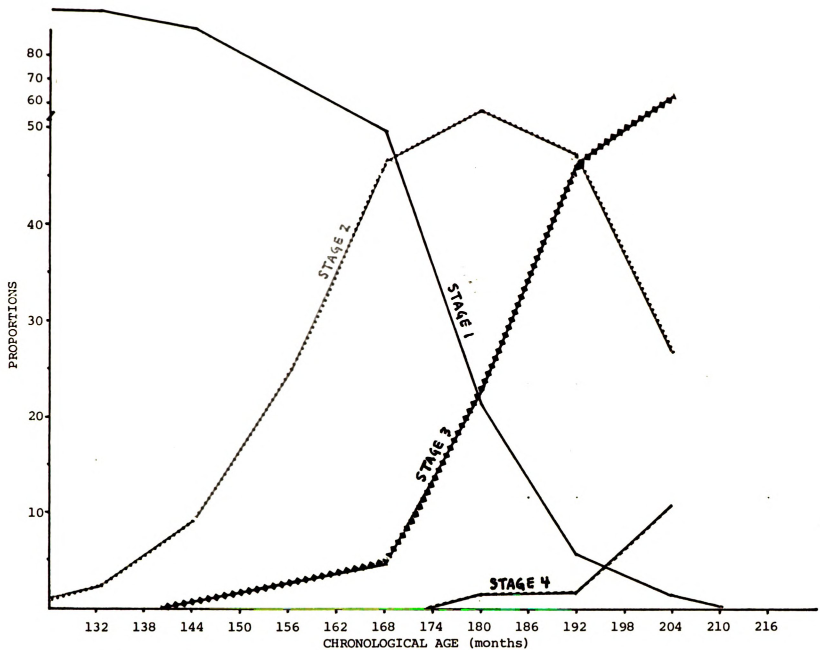


Figure 2.9 Observed laryngial stage proportions (Neyzi et al, 1975)

characteristics of penis, and testes and scrotum develop in a manner so orderly that many researchers have thoroughly described the stages through which those characteristics pass in their progress toward maturity, and 3) the adolescent male's secondary sexual characteristics of facial hair, axillary hair, and pubic hair also develop in a manner so orderly that many researchers have thoroughly described the stages through which those secondary sexual characteristics pass in their progress toward maturity.

RELATIONSHIP BETWEEN MOTOR PROFICIENCIES AND VARIOUS CLASSIFICATION PROCEDURES FOR ADOLESCENT MALES

The purpose of athletic competition is to reward the competitors putting forth the greatest effort and those with the higher athletic ability. Therefore, to prevent a biasing effect that favors those adolescent males with early onset and accelerated rates of biological maturation, sponsors of adolescent male competitions should equate the boys according to biological age. For sixty years, specialists in child growth and development have attempted to equate athletic competitions for adolescent males by using various indices to classify them.

To determine an administratively feasible index by which primary and secondary schools could homogeneously classify their boys for physical education classes, Indiana University Instructor Karl W. Bookwalter intensively surveyed the literature on various classification indices (Bookwalter, 1939). After examining McCloy's Classification Index, a Force Index, Cozen's General Athletic Ability Index, and Individual Athletic Events Index, Brace's Motor Ability Index, MacCurdey's Physical Capacity Index, Roger's Physical Fitness Index, and Roger's Strength Index, Bookwalter concluded that elementary schools should classify boys for athletic competition either by Roger's Strength Index or McCloy's Classification Index, and Junior High Schools should classify boys either by Roger's Strength Index or by MacCurdy's Physical Capacity Index.

To study the relationship between motor performance and the classification indices of skeletal age, physiological maturity, and the anthropometric measurements of height, weight, stem length/height, and breadth/length, University of California Institute of Child Welfare worker Anna S. Espenschade studied adolescent boys from the University of California's Adolescent Study (Espenschade, 1940). Using skeletal age values assessed by Professor Nancy Bayley, Espenschade found positive correlations between the skeletal maturity and motor performances of adolescent males. Using Professor Stolz's three sexual development phases for adolescent males, Espenschade found a direct relationship between motor progression and sexual development. In several of the motor performance events, she found that the motor performance and weight advanced with increasing age. She found that whereas the relationship between weight and ability to throw a ball for distance decreased somewhat over time, the relationship between weight and standing long jump scores increased steadily over time.

To contrast maturational, structural, strength, and motor traits of fifth through ninth grade boys, University of Oregon Professor H. Harrison Clarke and Kay H. Peterson studied 202 boys from the 1956-1957 Medford, Oregon, Boy's Growth Study (Clarke and Peterson, 1961a). They divided the boys into athletic participants and non-participant groups. Within the athletic participant group, they asked the coaches to rate the boys as: 1) exceptional athletes, 2) good athletes, 3) "out for the sport" athletes. Clarke and Peterson did not describe how they determined skeletal age. Nevertheless, they found that the outstanding elementary school athletes were more skeletally mature than their lesser skilled and non-participating peers. They found that the outstanding elementary school athletes exceeded their lesser skilled and non-participating peers in height, weight, and lung capacity measurements. They found that the outstanding junior high school athletes also had significantly higher mean skeletal ages than their lesser skilled and non-participating peers. And, in addition to greater mean values for height, weight, and lung capacity, they found that the outstanding junior high school athletes also exceeded their lesser skilled and non-participating peers in arm girth, chest girth, and calf girth measurements. In their conclusion, Clarke and Peterson stated the outstanding elementary and junior high school athletes were definitely more skeletally mature, taller and heavier, and stronger than their lesser skilled and non-participating peers.

To examine the differences in physical and motor traits of boys with advanced, normal, and retarded skeletal maturity, University of Oregon Professor H. Harrison Clarke and James C. E. Harrison studied 273 boys from the 1956-1957 Medford, Oregon, Boys' Growth Study aged 9, 12, and 15 years old (Clarke and Harrison, 1962a). As their measure of biological maturity they assessed hand-wrist X-rays by a Clarke-Hayman short method. For the three maturity groups at all three ages, they found the greatest mean differences in average body weight. When they compared the three groups on the basis of grip strength, back lift strength, leg lift strength, pushup ability, pullup ability, Roger's arm strength, Roger's Strength Index, Roger's Physical Fitness Index, mean cable-tension strength, elbow flexion strength, shoulder flexion strength, shoulder inward rotation strength, and standing long jump ability, the only variables for which they did not find significant differences in favor of the boys with advanced skeletal maturation over those with retarded skeletal maturation were pushups and pullups. However, because they knew that both body weight and muscular strength increased with advancing skeletal maturity, thus equalizing each variable's gain, they concluded that they should have expected the pushups and pullups results.

To construct maturity, structural, strength, and motor ability development curves for adolescent boys, University of Oregon Professor H. Harrison Clarke and J. Stuart Wickens studied 280 boys from the 1956-1957 Medford, Oregon, Boys' Growth Study. There were forty boys per group at the seven annual ages from 9 through 15 years old (Clarke and Wickens, 1962b). While they did not precisely describe the method that they used to assess skeletal age, they did reference the Greulich and Pyle hand-wrist Atlas. They reported the average skeletal age at each annual age for their sample (Table 2.120).

By classifying their sample in annual chronological age groups, they found that the mean 60-yard shuttle run development curve plateaued

Table 2.120

The skeletal age mean values and standard deviations for the boys across their chronological ages (Clarke and Wickens, 1962b)

Descriptive Data	Chronological Age						
	9	10	11	12	13	14	15
Mean (mo.)	106.12	114.43	131.75	146.25	160.52	173.62	184.25
S.D. (mo.)	1.82	1.89	2.08	1.91	2.04	1.77	1.21

between their 9 to 10 year old boys, rose sharply between their 10 to 12 year old boys, dipped between their 12 to 13 year old boys, rose sharply between their 13 to 14 year old boys, and plateaued between their 14 to 15 year old boys. Moreover, the mean standing long jump development curve had a nearly straight line rise between their 9 to 12 year old boys, dipped between their 12 to 13 year old boys, rose slightly between their 13 to 14 year old boys, and decreased between their 14 to 15 year old boys.

To determine the relationship between reaction time, movement time, and task completion time to adolescent boys' motor performances, strength scores, anthropometric measurements, and pubescent sexual development stages, University of Oregon Professor H. Harrison Clarke and Don Glines studied data on 65 boys within two months of their thirteenth birthday (Clarke and Glines, 1962c). As their measure of sexual development, they used Greulich's 1942 pubic hair and genitalia development indices. They found that their estimate of the boys' sexual development had no significant correlation with their total body reaction, arm reaction, movement time, or task completion times.

To compare adolescent males' skeletal age, various anthropometric measurements, and various motor performance abilities with their sexual development status, University of Oregon Professor H. Harrison Clarke and Earnest W. Degutis studied 236 boys from the 1956-1957 Medford, Oregon, Boys' Growth Study aged 10, 13, and 16 years old (Clarke and Degutis, 1962d). They used the Clarke-Hayman short method for their skeletal age measure. By using the criteria proposed by Greulich (1942), they assessed sexual development. They reported the observed frequency of boys in Greulich's five sexual maturation stages (Table 2.121). In parentheses, the author added the proportions that represented the frequency at each chronologic age.

Table 2.121

The frequency of the boys in Greulich's five Maturity Groups (Clarke and Degutis, 1962d)

		Greulich Maturity Groups				
Ages	N	1	2	3	4	5
10	86	72(83.7)	14(16.3)	0(0.0)	0(0.0)	0(0.0)
13	65	2(3.1)	25(38.5)	26(40.0)	8(12.3)	4(6.2)
16	86	0(0.0)	0(0.0)	0(0.0)	12(14.0)	78(86.0)

For 10, 13, and 16 year old boys, they found that when sexual development advanced, skeletal ages also advanced. For their 13 year old boys, they found that the skeletal age of the Maturity Group 2 boys was six months delayed from their average chronological age, that the skeletal age of the Maturity Group 3 boys was more than three months advanced over their average chronological age, and that the skeletal ages of the Maturity Groups 4 and 5 boys were more than fifteen and one-half months advanced over their average chronological age. The data for the 16 year old boys revealed that the skeletal ages of the Maturity Group 4 boys were twelve months delayed from their average chronological age, and those of Maturity Group 5 boys were seven months advanced over the average chronological age. Clarke and Degutis concluded that although Greulich's Maturity Group Index was not as sensitive to maturation changes as skeletal age for the pre-pubescent 10 and post-pubescent 16 year old boys, it was effective for the pubescent 13 year old boys.

To determine the value of age, height, and weight as classification indexes for secondary school students performing the AAHPER's Youth Fitness tests, Pennsylvania State University Professors Elmer A. Gross and Jerome A. Casciani analyzed data that physical education teachers in 82 Pennsylvanian secondary schools collected on over 13,000 junior and senior high school boys and girls (Gross and Casciani, 1962). They found that with the exception of junior high school boys performing the softball throw test, the standing long jump test, and the 50-yard dash, whether singly or in combination, age, height, and weight measures had little value in classifying students for motor proficiency tests.

To determine the relationships between age, height, and weight to motor performances of adolescent boys, University of California Professor Anna S. Espenschade studied the 50-yard dash, standing long jump, throw for distance, situp, and pullup performances of 3,800 California boys aged 10 to 18 years old (Espenschade, 1963). By placing the 12 to 15 year old boys into one group, Espenschade found that age, height, and weight accounted for twenty-five percent of the variance in their 50-yard dash, standing long jump, and throw for distance performances. With the single exception of situps, she found that all test performances increased steadily at all annual age increments.

Except to say that undoubtedly physical size was more similar within its groupings, Espenschade did not elaborate on the California Classification Plan. However, she reported that when she grouped boys according to physical size, she found the California Classification Plan superior to chronological age alone. Nevertheless, because 1) the California Classification Plan evolved thirty years ago and 2) today's boys are taller and heavier, she advised that someone update the California Classification Plan's grouping criteria.

Although she found that an Index created by combining the boys' age, height, and weight predicted the performances of a few adolescent boys more accurately than their age alone, Espenschade believed that the amount of prediction improvement was not sufficient to justify the required added time and effort. Therefore, Espenschade recommended that researchers use chronological age alone as their basis for developing motor performance test norms.

To examine the stability of dynometric muscular strength and motor performance ability from childhood to adolescence and to determine the predictive ability of motor performances at adolescence as predicted from their childhood body size, muscular strength, and motor performance measurements, University of Wisconsin Professors G. Lawrence Rarick and Frank L. Smoll annually measured 25 boys from age 7 to 12 years old (Rarick and Smoll, 1967). To obtain their adolescent values, they measured the boys a seventh time five years later. Their anthropometric data included height and weight from which they determined the reciprocal ponderal index. Their dynometric muscular strength measure included wrist flexion, elbow flexion, shoulder abduction, shoulder medial rotation, ankle extension, knee extension, hip flexion, and hip extension. Measures of gross motor skills included the 30-yard dash, overarm throw for velocity, and standing long jump.

Rarick and Small found that from ages 7 to 12 years and from childhood to age 17, height, weight, physique, standing long jump, and 30-yard dash developments showed relatively stable growth trends. However, upper extremity dynometric muscular strength and the overarm throw for velocity developed in an unstable manner. The best combination of physical growth, muscular strength, and motor performance measurements that they took on the boys when they were 10 years old accounted for seventy-eight percent of the standing long jump scores and fifty-nine percent of the 30-yard dash scores of the boys when they reached 17 years old. They concluded that boys with good performances in selected motor abilities during their childhood did not necessarily retain those performance levels during adolescence. However, if boys were strong and skilled in running and jumping during their childhood, then they were more likely to rank high in those skills during adolescence.

In 1971 University of Oregon Physical Education Research Professor H. Harrison Clarke compiled all of the studies that he and his graduate assistants had reported on the 1956 to 1968 Medford, Oregon, Boys' Growth Study and published it in a book entitled *Physical and Motor Tests in the Medford Boys Growth Study* (Clarke, 1971). From their extensive research Clarke drew the following conclusions: 1) exceptional adolescent male athletes are more skeletally mature, taller, heavier, stronger, more powerful, and more highly skilled than their peers; 2) although the characteristics differ at the three school levels, some fundamental characteristics of 12 year old elementary school athletes, 15 year old junior high school athletes, and 17 year old senior high school athletes evidenced themselves when the boys were younger; 3) the longitudinal significance of exceptional athletes' physical characteristics may extend to specific sports, and for various sports and at different school levels, those characteristics may vary; 4) as measured by their relative strength, especially for elementary and junior high school boys, general physical fitness indirectly contributed to a higher physical fitness of the athletes; 5) at all school levels, athletes who participated in all interscholastic sports scored consistently high in the gross strength measures; 6) at all school levels, athletes who participated in all interscholastic sports scored consistently high in the explosive power measures; 7) although successful athletes generally possessed common characteristics, the quality of each of those characteristics varied between athletes, such that, if athletes were weak in some characteristics, then by being very strong in other characteristics they compensated for weak characteristics; and 8) outstanding elementary school athletes did not necessarily become the outstanding junior high school athletes, and the outstanding junior high school athletes did not necessarily become the outstanding senior high school athletes.

Summary

From a review of the research literature, some of the principles concerning the relationships between motor proficiency levels in adolescent males and various classification indices pertinent to his investigation were: 1) biologically advanced adolescent males generally possess an advantage in motor performance over their biologically normal and delayed peers; 2) researchers have not found a classification index that is more desirable than chronological age from which they can predict the motor proficiencies of adolescent males on whom they have collected cross-sectional data; and 3) to identify a classification index that predicts the future motor proficiency of adolescent males requires longitudinal data.

The review of literature contained in Chapter II encompassed the research histories of the assessment of hand-wrist skeletal age, the assessment of sexual maturation in adolescent males, and the relationship between the motor proficiency ability of adolescent males and various classification procedures. In the review of the assessment of hand-wrist skeletal age, specialists demonstrated that estimates of skeletal age from hand-wrist X-rays represented the best predictor of biological age that researchers in child growth and development can use. Moreover, the X-ray standards contained in the Greulich and Pyle hand-wrist Atlas are reliable guides for the assessment of skeletal maturation.

In the review of the assessment of sexual maturation of adolescent males, the specialists showed that raters can reliably assess the sexual characteristics of facial hair, axillary hair, pubic hair, penis development, and scrotum and testes development. The estimate of sexual maturation represented the second best predictor of biological age in adolescent males that is currently available.

The review of the relationship between the motor proficiency ability of adolescent males and various classification procedures demonstrated that 1) the use of chronological age does not take into account the vast differences in the biological maturity of adolescent males and that 2) researchers need to collect longitudinal data to investigate predicting future motor proficiency.

The review of the literature provided the following findings that are pertinent to this investigation: 1) hand-wrist X-ray skeletal age estimates are the best estimates of the biological age of adolescent males; 2) the X-ray standards contained in the Greulich and Pyle hand-wrist Atlas are reliable and valid guides for assessing skeletal age; 3) the primary and secondary sex characteristics of facial hair, acillary hair, pubic hair, penis development, and scrotum and testes development provide reliable criteria for the sexual maturation of adolescent males; and 4) specialists in classifying adolescent males for motor proficiency have tried but rejected the use of biological maturation as a means of classifying participants for athletic competition.

CHAPTER III

METHODS AND PROCEDURES

Because finding answers to the four questions in this desserta-tion required two samples, chapter three provides two explanations of the experimental designs, data collections, and measures taken.

Experimental Design #1

Questions one, two, and three needed a sample of adolescent males who represented a rectangular distribution of sexual maturation. In order to randomly select such a sample, it was necessary to know the sexual maturation level of the potential subjects. Therefore, in conjunction with the Fall AAHPER Youth Fitness testing, the boys physical education teachers at East Lansing Michigan's Macdonald Middle School used the SMIV checklist to assess the sexual maturation of their students. Also, the author assessed the sexual maturation of the ninth grade boys in East Lansing's High School. A letter requesting permission for their son to participate in this study was sent to the parents. Forty-eight percent of the parents returned those forms and forty-seven percent were affirm-ative. From groupings based on sexual maturation levels, 30 subjects were randomly selected. Table 3.1 provides the SMIV score, chronological age, and grade of the boys.

Data Collection #1

All of the data were collected during a single Saturday morning session.

Table 3.1

The SMIV scores, chronological age, and grade levels of the 30 boys selected to represent a rectangular distribution of sexual maturation

Subject	SMIV	Chronological Age (months)	Grade
JB	1.1	134.5	7
JH	1.2	150.0	6
FF	1.3	145.5	6
BM	1.6	137.5	6
HS	1.7	143.0	7
RS	1.8	155.5	7
TB	1.9	143.5	6
JC	2.0	152.5	8
DD	2.1	170.0	9
ST	2.2	173.0	9
CM	2.3	179.0	9
BO	2.6	150.0	6
DF	2.7	138.0	6
JS	2.8	146.0	6
KB	2.9	163.0	8
JP	3.0	161.5	8
JL	3.1	170.5	9
TS	3.2	156.5	8
DB	3.3	173.0	9
CG	3.6	173.0	9
JP	3.7	153.0	7
CS	3.8	174.0	9
PW	3.9	164.0	8
PS	4.0	162.0	8
PZ	4.1	169.5	9
RH	4.2	154.0	7
DT	4.3	160.0	7
CO	4.6	175.5	8
DK	4.7	169.5	9
DP	4.8	180.0	8

Measures #1

Skeletal Age

Michigan State University Olin Health Center Radiologists X-rayed both hands and wrists of the subjects. (Appendix B contains a detailed description of the methodology.)

Using the Greulich-Pyle hand-wrist atlas, Michigan State University Professor Vernal D. Seefeldt viewed hand-wrist X-rays of the subjects and on two separate occasions determined the skeletal age (Appendix C). For each subject's skeletal age, the computer averaged the two ratings.

Physical Growth

Michigan State University Olin Health Center Radiologists measured the height and weight of the subjects. They recorded height to the nearest one-quarter inch and the analysis rounded .25 to .3 and .75 to .8. They recorded weight to the nearest one-half pound (Appendix D).

Sexual Maturation

Using the SMIV checklist, two East Lansing, Michigan and two Mason, Michigan boys physical education teachers and the author assessed sexual maturation. Where the data collection required only a single score for the five sexual characteristics, the analysis averaged the ten scores assigned each trait (Appendix E).

Chronological Age

School records provided the age of the subjects on the date of the data collection. Their age was recorded to the nearest one-half month.

Experimental Design #2

Question 4 needed a large, cross-sectional sample of adolescent males.

Data Collection #2

Beginning with the second week in April, every six months for three years, the East Lansing teachers assessed the seventh, and eighth grade male populations and the Mason teachers assessed the seventh, eighth, and ninth grade male population. Additionally, in the final test period, the East Lansing teachers assessed the sixth grade boys in East Lansing's new Macdonald Middle School. For those subjects on whom the teachers collected data for more than one test period, because the sample needed more subjects at higher SMIV scores, the analysis included their last complete data set. Table 3.2 contains the number of East Lansing subjects in each grade for each test period. Of the 878 East Lansing boys in this

Table 3.2

The number of East Lansing subjects in each grade for each test period

TEST	GRADE LEVEL				
PERIOD	SIXTH	SEVENTH	EIGHTH	NINTH	
1	0	48	76	0	
2	0	68	156	0	
3	0	41	48	0	
4	0	49	57	0	
5	0	74	40	0	
6	83	96	42	0	
TOTALS	83	376	419	0	878

study, 83 were sixth graders, 373 were seventh graders, and 418 were eighth graders. Table 3.3 contains the number of Mason subjects in each grade for each test period. Of the 700 Mason boys in this study, 207

Table 3.3

The number of Mason subjects in each grade for each test period.

TEST	GRADE LEVEL				
PERIOD	SIXTH	SEVENTH	EIGHTH	NINTH	
1	0	31	30	12	
2	0	29	26	11	
3	0	28	36	115	
4	0	31	32	19	
5	0	60	51	89	
6	0	28	30	42	
TOTALS	0	207	205	288	700

were seventh graders, 205 were eighth graders, and 288 were ninth graders.

Measures #2

School

The boys attended either East Lansing or Mason.

Test Period

Test period was recorded as; first, second, third, fourth, fifth, or sixth.

Grade

Grade was recorded as; sixth grade in the fall, seventh grade in the fall, seventh grade in the spring, eighth grade in the fall, eighth grade in the spring, ninth grade in the fall, or ninth grade in the spring.

Chronological Age

School records provided the age of the subjects on the date of the data collection. Age was recorded to the nearest one-half month.

Physical Growth

At each school, the teachers measured the height and weight of the boys. They recorded height to the nearest one-quarter inch and the analysis rounded .25 to .3 and .75 to .8. They recorded weight to the nearest one-half pound (Appendix F).

AAHPER Youth Fitness Tests

The teachers tested the following motor proficiency skills; 1) sit-ups with a maximum of 99, 2) pull-ups, 3) standing long jump to the nearest inch, 4) vertical jump, to the nearest inch, and 5) shuttle run, to the nearest one-tenth of a second (Appendix G).

Sexual Maturation

Using the SMIV checklist, the teachers assessed the sexual maturation of the boys. For each sexual characteristic, the analysis used the number of the stage that the teachers checked (Appendix H).

Treatment of the Data

a) Question one: What proportion of the variance in the skeletal age of adolescent males is explained by increments in their chronological age, height, and weight?

The Multivariate Analysis of Variance computer program determined the square of the multiple r values and the step-wise regression analysis. The square of the multiple r provided the proportion of the dependent variable (skeletal age) that is explained by the increments in the independent variables (chronological age, height, weight, and SMIV). The step-wise regression analysis explained the contribution to the variance in the dependent variable that is unique to each independent variable. Therefore, question one became a five-step statistical analysis.

1) What is the square of the multiple r for estimating skeletal age from chronological age?

2) What is the square of the multiple r value for estimating skeletal age from chronological age and height?

3) Was the additional contribution unique to height statistically significant?

4) What is the square of the multiple r value for estimating skeletal age from chronological age, height, and weight?

5) Was the additional contribution unique to weight statistically significant?

Educational researchers typically choose the .05 significance level. Therefore, in this study if the analysis falsely accepts that height and weight do contribute significantly, then the estimate of skeletal age (ESA) will incorrectly include those variables. However, if the analysis falsely rejects the significant contributions of height and weight, then ESA will incorrectly exclude those variables.

b) Question two: Does SMIV explain a significant additional proportion of the variance in skeletal age? By using the same statistical procedures as in question one, question two became a two-step statistical analysis.

1) What is the square of the multiple r value for estimating skeletal age from chronological age, height, weight, and SMIV?

2) Was the additional contribution unique to SMIV statistically significant?

If the analysis falsely accepts that SMIV does contribute significantly, then ESA will incorrectly include the SMIV variable. However, if the analysis falsely rejects the significant contribution of SMIV, then ESA will incorrectly exclude the SMIV variable.

c) Question three: Can Boys Physical Education teachers reliably assess the sexual maturation of adolescent males?

The correlation analysis determined the within-rater reliabilities and the between-rater reliabilities. However, between-rater correlations combine the errors between raters, across sexual characteristics, and over time which results in an underestimation of the between-rater reliability. Therefore, an extension of the intraclass correlation method of estimating reliabilities that bases its procedure on analysis of variance calculated the individual reliabilities for between raters, across sexual characteristics, and over time.

By conceptualizing this analysis in the following two ways, the intraclass correlation method provided eight separate reliabilities; 1) researchers using the five sexual manifestations individually and equally as indicants of sexual maturation and 2) researchers using the average score of the five sexual manifestations as the indicant of sexual maturation.

1) ρ WITHIN, or for a fixed rater, a fixed sexual characteristic, and a fixed time point, what was the reliability coefficient?

2) ρ STABILITY OVER TIME, or for a fixed rater and a fixed sexual characteristic, what was the reliability coefficient?

3) ρ EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS, or for a fixed rater and a fixed time point, what was the reliability coefficient?

4) ρ STABILITY and EQUIVALENCE, or for a fixed rater, what was the reliability coefficient?

5) ρ RATERS, or for a fixed sexual characteristic and a fixed time point, what was the reliability coefficient?

6) ρ RATERS STABILITY, or for a fixed sexual characteristic, what was the reliability coefficient?

7) ρ RATERS EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS, or for a fixed time point, what was the reliability coefficient?

8) ρ RATERS STABILITY and EQUIVALENCE, or across raters, across sexual characteristics, and over time, what was the reliability coefficient?

The reliability coefficient that directly answers question three is the ρ RATERS when the analysis used the average of the five sexual manifestations as the indicant of sexual maturation.

d) Question four: How does an estimate of skeletal age (ESA) that encompasses chronological age, height, weight, and SMIV compare with chronological age when classifying adolescent males for motor proficiency norms?

The following statistics examined this question:

1) within cell variances

Smaller within-cell variances indicate more homogeneous groups.

2) mean value increments

Larger mean value differences across the annual classifications indicate more homogeneous groups.

CHAPTER IV

RESULTS AND DISCUSSION

Question I. What proportion of the variance in the skeletal age of adolescent males is explained by increments in their chronological age, height, and weight?

The Multivariate Analysis of Variance computer program statistically analyzed the data (Schiefley and Schmidt, 1973). The square of the multiple r estimated the percent of the variance in the dependent variable explained by the independent variables.

In the first part of question I, the dependent variable was skeletal age and the independent variable was chronological age. When chronological age was the only independent variable, the square of the multiple r was .440 or approximately 45 percent. Therefore, the proportion of the variance in skeletal age explained by chronological age was 45 percent.

In the second part of question I, while the dependent variable remained skeletal age, the independent variables became chronological age and height. When chronological age and height were the independent variables, the square of the multiple r was .648 or 65 percent. Therefore, the proportion of the variance in skeletal age explained by the contribution of chronological age and height was 65 percent. The inclusion of height uniquely explained an additional 20 percent of the variance.

As a result of part two, the third part of question I became; Was the additional 20 percent of the variance in skeletal age unique to height statistically significant? Step-wise regressions determine the significance levels of additional contributions by each independent variable. To interpret step-wise regressions, when the contribution of the last variable is significant, it is not possible to discuss the significance level of the preceding variable. However, when the contribution of the last variable is not significant, it is possible to discuss the significance level of the preceding variable.

Table 4.1 contains the step-wise regression for chronological age and height. Because the significance level for height was less than .05,

Table 4.1

The step-wise regression for chronological age and height

	Percent of Additional Variance	Significance Level
Chronological Age	44.937	
Height	19.903	P less than .0006

the additional contribution unique to height was significant.

In the fourth part of question I, while the dependent variable remained skeletal age, the independent variables became chronological age, height, and weight. When chronological age, height, and weight were the independent variables, the square of the multiple r was .708 or approximately 71 percent. Therefore the proportion of the variance in skeletal age explained by the combination of chronological age, height, and weight was 71 percent. The inclusion of weight uniquely explained an additional 6 percent of the variance.

As a result of part four, the fifth part of question I became; Was the additional 6 percent of the variance in skeletal age unique to weight statistically significant? Table 4.2 contains the step-wise regression for chronological age, height, and weight. Because the significance level for weight was less than .05, the additional contribution unique

Table 4.2

The step-wise regression for chronological age, height, and weight

	Percent of Additional Variance	Significance Level
Chronological Age	44.937	
Height	19.903	
Weight	5.932	P less than .0300

to weight was significant.

To summarize, the answers to question I's five sub-questions are;

1) What is the square of the multiple r for estimating skeletal age from chronological age? For this sample, that square of the multiple r value was .449.

2) What is the square of the multiple r for estimating skeletal age from chronological age and height? For this sample, that square multiple r value was .648.

3) Was the additional contribution unique to height statistically significant? For this sample, height did uniquely explain a significant additional proportion of the variance.

4) What is the square of the multiple r value for estimating skeletal age from chronological age, height, and weight? For this sample, that square of the multiple r value was .708.

5) Was the additional contribution unique to weight statistically significant? For this sample, weight did uniquely explain a significant additional proportion of the variance.

At this point some readers might ask, if the order of the independent variables reversed, then what would happen to the step-wise regressions (Appendix I)?

Question II. Does SMIV explain a significant additional proportion of the variance in skeletal age?

In the first part of question II, the dependent variable remained skeletal age while the independent variables included chronological age, height, weight, and the five sexual characteristics of; facial hair axillary hair, pubic hair, penis development, and scrotum and testes development. The square of the multiple r was .839 or approximately 84 percent. Therefore, the proportion of the variance in skeletal age explained by the combination of chronological age, height, weight, and the five sexual characteristics was 84 percent. The inclusion of the five sexual characteristics variables uniquely explained an additional 13 percent of the variance.

Because doing a step-wise regression analysis with chronological age, height, weight, and the five sexual characteristics would give us the significance levels for the five sexual characteristics individually, a SMIV regression equation that combined the five sexual characteristics was determined (Appendix J).

When the independent variables were chronological age, height, weight, and SMIV, the square of the multiple r was .828 or approximately 83 percent. The inclusion of SMIV uniquely explained an additional 12 percent of the variance.

As a result of part one, the second part of question II became; Was the additional 12 percent of the variance in skeletal age unique to SMIV statistically significant? Table 4.32 contains the step-wise regressions for chronological age, height, weight, and SMIV. Because the significance level for SMIV was less than .05, the additional contribution unique to SMIV was significant.

To summarize, the answers to question II's two sub-questions are;

1) What is the square of the multiple r value for estimating

Table 4.3

The step-wise regression for chronological age, height, weight and SMIV

	Percent of Additional Variance	Significance Levels
Chronological Age	44.937	
Height	19.850	
Weight	5.966	
SMIV	12.014	P less than .0004

skeletal age from chronological age, height, weight, and SMIV? For this sample, that square of the multiple r value was .828.

2) Was the additional contribution unique to SMIV statistically significant? For this sample, SMIV did uniquely explain a significant additional proportion of the variance.

At this point some readers might again ask, if the order of the independent variables reversed, then what would happen to the step-wise regressions (Appendix K)?

Question III. Can Boys Physical Education teachers reliably assess the sexual maturation of adolescent males?

Correlation coefficients estimated the within rater reliability between the first and second assessments of the five raters on each of the five sexual characteristics (Table 4.4).

Table 4.4

The within-rater correlation coefficients for the five raters across the five sexual characteristics and the last column shows the mean correlation across raters for each variable

Sexual Characteristics	RATERS 1	2	3	4	5	Average across Raters
Facial hair	.789	.683	.949	.492	.649	.772
Axillary hair	.777	.915	.926	.694	.877	.838
Pubic hair	.762	.940	.838	.883	.904	.865
Penis	.874	.931	.954	.851	.885	.899
Scrotum and Testes	.859	.916	.818	.710	.827	.826

The within-rater correlations for facial hair ranged from .949 to .492 while the last column shows that the mean correlation across raters was .77. For axillary hair, the within-rater correlations ranged from .962 to .694 and the average correlation across raters was .838. For pubic hair, the within-rater correlations ranged from .940 to .762 with the .865 averaged correlation across raters. For penis development, the within-rater correlations ranged from .954 to .851 with an average across raters correlation of .899. And, for scrotum and testes development, the within-rater correlations ranged from .916 to .710 with an average correlation across raters of .826. From highest to lowest, the order of the average correlation across raters was; 1) penis development, 2) pubic hair, 3) axillary hair, 4) scrotum and testes development, and 5) facial hair.

Because correlation statistics describe how well two values approximate each other, those within-rater correlations told us only how well each rater agreed with himself. However, just because a rater's correlation shows that he agreed highly with himself, that did not necessarily indicate that that rater correctly assessed the boys' sexual characteristics. To determine how well the raters agreed with each other, each rater was paired with every other rater, and then those raters' first and second assessments were correlated (Tables 4.5 through 4.9). Table 4.10 contains each raters' average paired betwee-raters correlations for the five sexual characteristics. And, Table 4.11 contains the average of all of the paired between raters correlations for the five sexual characteristics.

Table 4.5

Between-raters correlations of facial hair for the paired five raters

Facial Hair	Raters									
Ratings	1/2	1/3	1/4	1/5	2/3	2/4	2/5	3/4	3/5	4/5
1/1	.618	.778	.490	.633	.576	.458	.578	.307	.612	.462
1/2	.734	.796	.706	.795	.728	.542	.627	.515	.681	.556
2/1	.724	.851	.443	.629	.698	.365	.620	.337	.690	.608
2/2	.710	.895	.577	.760	.781	.667	.738	.622	.760	.633
Average	.696	.830	.554	.704	.696	.508	.591	.445	.696	.565

Table 4.6

Between-raters correlations of axillary hair for the paired five raters

Axillary Hair	Raters									
Ratings	1/2	1/3	1/4	1/5	2/3	2/4	2/5	3/4	3/5	4/5
1/1	.733	.734	.716	.770	.829	.811	.809	.714	.851	.907
1/2	.721	.653	.716	.785	.787	.827	.862	.781	.853	.796
2/1	.843	.737	.744	.770	.828	.787	.786	.705	.853	.747
2/2	.848	.740	.734	.816	.821	.802	.821	.723	.840	.800
Average	.786	.716	.728	.785	.816	.807	.820	.731	.849	.813

Table 4.7

Between-raters correlations of pubic hair for the paired five raters

Pubic hair	Raters									
Ratings	1/2	1/3	1/4	1/5	2/3	2/4	2/5	3/4	3/5	4/5
1/1	.797	.756	.818	.819	.795	.913	.935	.866	.861	.890
1/2	.833	.763	.800	.780	.917	.861	.891	.828	1748	.848
2/1	.837	.763	.767	.875	.830	.886	.945	.869	.874	.854
2/2	.895	.824	.797	.833	.909	.874	.882	.805	.784	.911
Average	.841	.771	.796	.827	.863	.884	.913	.842	.817	.876

Table 4.8

Between-raters correlations of penis development for the paired five raters

Penis dev.	Raters									
Ratings	1/2	1/3	1/4	1/5	2/3	2/4	2/5	3/4	3/5	4/5
1/1	.904	.828	.782	.765	.925	.904	.894	.876	.880	.841
1/2	.876	.818	.761	.861	.890	.815	.901	.820	.837	.859
2/1	.914	.877	.920	.840	.871	.909	.843	.824	.883	.755
2/2	.932	.823	.823	.914	.840	.827	.862	.796	.815	.811
Average	.907	.837	.824	.845	.882	.889	.875	.829	.854	.817

Table 4.9

Between-raters correlations of scrotum and testes development for the paired five raters

Scrotum dev. Ratings	Raters									
	1/2	1/3	1/4	1/5	2/3	2/4	2/5	3/4	3/5	4/5
1/1	.854	.516	.701	.682	.603	.826	.801	.621	.571	.760
1/2	.835	.647	.764	.755	.721	.800	.889	.403	.571	.792
2/1	.881	.516	.811	.761	.671	.875	.792	.688	.651	.711
2/2	.887	.728	.764	.787	.741	.813	.840	.528	.627	.821
Average	.864	.602	.760	.746	.687	.836	.831	.556	.605	.776

Table 4.10

Average between-raters correlations of the paired raters for the five sexual characteristics

Sexual Characteristics	Raters					
	1	2	3	4	5	Ave.
Facial hair	.696	.623	.667	.518	.639	.629
Axillary hair	.754	.807	.778	.770	.817	.785
Pubic hair	.809	.875	.823	.850	.858	.843
Penis dev.	.853	.888	.851	.840	.848	.856
Scrotum dev.	.743	.805	.613	.732	.740	.727
Average	.771	.800	.746	.742	.780	.768

Table 4.11

Average of all of the paired between-raters correlations for the five sexual characteristics

Sexual Characteristics				
Facial Hair	Axillary Hair	Pubic Hair	Penis Development	Scrotum Development
.629	.785	.843	.856	.727

For facial hair, rater 1's .696 average paired between-raters correlation was the highest of the five raters. The .692 average of all of the paired between-raters correlations placed facial hair fifth in this list of five characteristics.

For axillary hair, rater 5's .817 average paired between-raters correlation was the highest and rater 1's .754 average correlation was the lowest. The average of all of the paired between-raters correlations of .785 placed axillary hair third.

For pubic hair, rater 2's .875 average paired between-raters correlation was the highest and rater 1's .809 average correlation was the lowest. The average of all of the paired between-raters correlations of .843 placed pubic hair second.

For penis development, rater 2's .888 average paired between-raters correlation was the highest and rater 4's .840 average paired between-raters correlation was the lowest. The average of all of the paired between-raters correlations of .856 placed penis development first.

For scrotum and testes development, rater 2's .805 average paired between-raters correlation was the highest and rater 3's .613 was the lowest. The average of all of the paired between-raters correlation of .727 placed scrotum and testes development fourth.

The averages of the average paired between-raters correlations for the five sexual characteristics contained in Table 4.11 demonstrate the consistency of the five raters. From highest to lowest ability of those five raters, the order of the raters was; rater 2 (.800), rater 5 (.750), rater 1 (.771), rater 3 (.746), and rater 4 (.742). The averages of all of the paired between-raters correlations suggest the agreement of the five raters. The order of the averages of all of the paired between-raters correlations for the five sexual characteristics were; penis development (.856), pubic hair (.843), axillary hair (.785), scrotum and testes development (.727), and facial hair (.629).

As the author's doctoral guidance committee's statistics advisor, Michigan State University College of Education Professor William H. Schmidt provided the following extension of the intraclass correlation method of estimating reliabilities based on analysis of variance techniques. The four variables in this analysis were; 1) Raters (R), 2) Sexual Characteristics (C), 3) Time points (T), and 4) Subjects (S). Because the Raters, Time points, and Subjects were random variables and the combined Sexual Characteristics was a fixed variable, this 4-way ANOVA analysis became a mixed model. By using Time as a variable rather than within-cell replications, each cell contained only one value.

As its first step, this analysis required the determination of the expected Mean Square (MS) formulas for a 4-way Analysis of Variance mixed model. By using the Millman-Glass Rules of Thumb for Writing the Anova Table, the Mean Square formulas contained in Table 4.12 were determined (Millman and Glass, 1967).

By assuming that those four variables did not interact, the σ^2(RCTS) value became the error estimate (σ^2(e)). Using the Table 4.12 Mean Square formulas, the formulas for the estimate of the variance components were isolated (Table 4.13).

After determining those formulas, an N-way analysis of variance computer program calculated the Mean Square values (Table 4.14) (Edon, 1976). By substituting those Mean Square values into the estimates of the variance components formulas, the estimates of the variance components were calculated (Table 4.15).

For the first series of reliabilities, the analysis assumed that the five primary and secondary sexual characteristics were manifestations of the underlying trait of sexual maturation. Therefore, those reliabil-

Table 4.12

The Mean Squares formulas for a 4-way Anova mixed model where Raters(R), Time points(T), and Subjects(S) are random and Sexual Characteristics(C) is fixed

Mean Squares			Formula
1.	RCTS	=	σ^2(RCTS)
2.	CTS	=	σ^2(RCTS) + $r\sigma^2$(CTS)
3.	RTS	=	$c\sigma^2$(RTS)
4.	RCS	=	σ^2(RCTS) + $t\sigma^2$(RCS)
5.	RCT	=	σ^2(RCTS) + $s\sigma^2$(RCT)
6.	TS	=	$c\sigma^2$(RTS) + $rc\sigma^2$(TS)
7.	CS	=	σ^2(RCTS) + $r\sigma^2$(CTS) + $t\sigma^2$(RCS) + $rt\sigma^2$(CS)
8.	CT	=	σ^2(RCTS) + $r\sigma^2$(CTS) + $s\sigma^2$(RCT) + $rs\sigma^2$(CT)
9.	RS	=	$c\sigma^2$(RTS) + $ct\sigma^2$(RS)
10.	RT	=	$c\sigma^2$(RTS) + $cs\sigma^2$(RT)
11.	RC	=	$c\sigma^2$(RTS) + $cs\sigma^2$(RT)
12.	S	=	$c\sigma^2$(RTS) + $cs\sigma^2$(RT)
13.	T	=	$c\sigma^2$(RTS) + $rc\sigma^2$(TS) + $cs\sigma^2$(RT) + $rcs\sigma^2$(T)
14.	C	=	σ^2(RCTS) + $r\sigma^2$(CTS) + $t\sigma^2$(RCS) + $s\sigma^2$(RCT) + $rt\sigma^2$(CS) + $rs\sigma^2$(CT)
15.	R	=	$c\sigma^2$(RTS) + $ct\sigma^2$(RS) + $cs\sigma^2$(RT) + $cts\sigma^2$(R)

Table 4.13

The estimates of the variance components formulas in this 4-way mixed model

Estimate of the Variance Components			Formulas
1.	$\hat{\sigma}^2$(e)	=	MS(RCTS)
2.	$\hat{\sigma}^2$(CTS)	=	MS(CTS) - MS(RCTS)/r
3.	$\hat{\sigma}^2$(RTS)	=	MS(RTS) - MS(RCTS)/c
4.	$\hat{\sigma}^2$(RCS)	=	MS(RCS) - MS(RCTS) t
5.	$\hat{\sigma}^2$(RCT)	=	MS(RCT) - MS(RCTS) s
6.	$\hat{\sigma}^2$(TS)	=	MS(TS) - MS(RCTS) - $c\hat{\sigma}^2$(RTS)/rc
7.	$\hat{\sigma}^2$(CS)	=	MS(CS) - MS(RCTS) - $t\hat{\sigma}^2$(RCS) - $r\hat{\sigma}^2$(CTS)/rt
8.	$\hat{\sigma}^2$(CT)	=	MS(CT) - MS(RCTS) - $s\hat{\sigma}^2$(RCT) - $r\hat{\sigma}^2$(CTS)/rs
9.	$\hat{\sigma}^2$(RS)	=	MS(RS) - MS(RCTS) - $c\hat{\sigma}^2$(RTS)/ct
10.	$\hat{\sigma}^2$(RT)	=	MS(RT) - MS(RCTS) - $c\hat{\sigma}^2$(RTS)/cs
11.	$\hat{\sigma}^2$(RC)	=	MS(RC) - MS(RCTS) - $s\hat{\sigma}^2$(RCT) - $t\hat{\sigma}^2$(RCS)/tx
12.	$\hat{\sigma}^2$(S)	=	MS(S) - MS(RCTS) - $ct\hat{\sigma}^2$(RS) - $rc\hat{\sigma}^2$(TS) - $c\hat{\sigma}^2$(RTS)/rct
13.	$\hat{\sigma}^2$(T)	=	MS(T) - MS(RCTS) - $cs\hat{\sigma}^2$(RT) - $rc\hat{\sigma}^2$(TS) - $c\hat{\sigma}^2$(RTS)/ rcs
14.	$\hat{\sigma}^2$(C)	=	MS(C) - MS(RCTS) - $ts\hat{\sigma}^2$(RC) - $rs\hat{\sigma}^2$(CT) - $rt\hat{\sigma}^2$(CS) - $s\hat{\sigma}^2$(RCT) - $t\hat{\sigma}^2$(RCS) - $r\hat{\sigma}^2$(CTS)/rts
15.	$\hat{\sigma}^2$(R)	=	MS(R) - MS(RCTS) - $cs\hat{\sigma}^2$(RT) - $ct\hat{\sigma}^2$(RS) - $c\hat{\sigma}^2$(RTS)/cts

Table 4.14

The Mean Square values for the fifteen sources of variance in this 4-way mixed model

Sources of Variance	Mean Squares
Raters-Sexual characteristics-Time-Subjects (RCTS)	.1798
Raters-Sexual characteristics-Subjects (RCS)	.3348
Raters-Time-Subjects (RTS)	.3476
Sexual characteristics - Time - Subjects (CTS)	.1903
Raters - Sexual characteristics - Time (RCT)	.1244
Raters - Subjects (RS)	.7641
Sexual characteristics - Subjects (CS)	2.4635
Time - Subjects (TS)	.4430
Raters - Sexual characteristics (RC)	6.6689
Raters - Time (RT)	.2190
Sexual characteristics - Time (CT)	.2323
Subjects (S)	44.7968
Raters (R)	16.3810
Sexual characteristics (C)	97.4477
Time (T)	.0540

Table 4.15

The substituted estimates of the variance components formulas

Estimate of the Variance Components		Final Values
1. $\hat{\sigma}^2(e)$	= .1798	.1798
2. $\hat{\sigma}^2(CTS)$	= .1903 - .1798/5	.0021
3. $\hat{\sigma}^2(RTS)$	= .3476 - .1798/5	.0036
4. $\hat{\sigma}^2(RCS)$	= .3348 - .1798/2	.0775
5. $\hat{\sigma}^2(RCT)$	= .1244 - .1798/30.	.0000
6. $\hat{\sigma}^2(TS)$	= .4430 - .1798 - .1678/5(5)	.0038
7. $\hat{\sigma}^2(CS)$	= .4635 - .1798 - .1550/5(2)	.2118
8. $\hat{\sigma}^2(CT)$	= .2323 - .1798 - .0000 - .0105 /5(30)	.0003
9. $\hat{\sigma}^2(RS)$	= .7641 - .1798 - .1678/5(2)	.0417
10. $\hat{\sigma}^2(RT)$	= .2190 - .1798 - .1678/5(30)	.0000
11. $\hat{\sigma}^2(RC)$	= 6.6689 - .1798 - .0000 - .1550/2(30)	.1056
12. $\hat{\sigma}^2(S)$	= 44.7968 - .1798 - .4165 - .0954 - .1678/5(5)(2)	.8787
13. $\hat{\sigma}^2(T)$	= .0540 - .1798 - .0000 - .0954 - .1678/5(5)(30)	.0000
14. $\hat{\sigma}^2(C)$	= .97.4470 - .1798 - 6.3895 - .0420 - 2.1182 - .0000 - .1550 - .0105/5(2)(30)	.2785
15. $\hat{\sigma}^2(R)$	= 16.3810 - .1798 - .0000 - .4165-.1678/5(2)(30)	.0521

ities represent researchers using any one of those five sexual characteristics as an indicant of sexual maturation (Table 4.16).

Table 4.16

The reliability formulas for the eight sources of error in this 4-way mixed model

Reliability	Formulas
1. ρ WITHIN	$\hat{\sigma}^2(S)/\hat{\sigma}^2(S)+\hat{\sigma}^2(e)$
2. ρ STABILITY OVER TIME	$\hat{\sigma}^2(S)/\hat{\sigma}^2(S)+\hat{\sigma}^2(TS)+\hat{\sigma}^2(e)$
3. ρ EQUIVALENCE ACROSS ρ SEXUAL CHARACTERISTICS	$\hat{\sigma}^2(S)/\hat{\sigma}\hat{\delta}(S)+\hat{\sigma}^2(CS)+\hat{\sigma}^2(e)$
4. ρ STABILITY and EQUIVALENCE	$\hat{\sigma}^2(S)/\hat{\sigma}^2(S)+\hat{\sigma}^2(TS)+\hat{\sigma}^2(CS)+\hat{\sigma}^2(e)$
5. ρ RATERS	$\hat{\sigma}^2(S)/\hat{\sigma}^2(S)+\hat{\sigma}^2(RS)+\hat{\sigma}^2(e)$
6. ρ RATERS STABILITY OVER TIME	$\hat{\sigma}^2(S)/\hat{\sigma}^2(S)+\hat{\sigma}^2(RS)+\hat{\sigma}^2(TS)+\hat{\sigma}^2(e)$
7. ρ RATERS EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS	$\hat{\sigma}^2(S)/\hat{\sigma}^2(S)+\hat{\sigma}^2(RS)+\hat{\sigma}^2(CS)+\hat{\sigma}^2(e)$
8. ρ RATERS STABILITY and EQUIVALENCE	$\hat{\sigma}^2(S)/\hat{\sigma}^2(S)+\hat{\sigma}^2(RS)+\hat{\sigma}^2(CS)+\hat{\sigma}^2(TS)+\hat{\sigma}^2(e)$

By substituting the calculated estimates of the variance components for their reliability formulas counterparts, the reliability coefficients were calculated (Table 4.17).

Table 4.17

The substituted reliability formulas and their final value

Reliability	Substituted formulas	Final Value
1.ρ WITHIN	.8787/.8787+.1798	.830
2.ρ STABILITY	.8787/.8787+.0038+.1798	.827
3.ρ EQUIVALENCE	.8787/.8787+.2118+.1798	.692
4.ρ STABILITY and EQUIVALENCE	.8787/.8787+.0038+.2118+.1798	.690
5.ρ RATERS	.8787/.8787+.0417+.1798	.799
6.ρ RATERS STABILITY	.8787/.8787+.0417+.0038+.1798	.794
7.ρ RATERS EQUIVALENCE	.8787/.8787+.0417+.2118+.1798	.670
8.ρ RATERS STABILITY and EQUIVALENCE	.8787/.8787+.0417+.0417+.0038+ .2118+.1798	.660

The ρ WITHIN indicated that for a fixed Rater, a fixed Sexual Characteristic, and a fixed Time point the reliability of a rating was .830. The ρ STABILITY OVER TIME indicated that for a fixed Rater and a fixed Sexual Characteristic, the reliability across time was .827. That finding suggests that raters can re-assess those five sexual characteristics with little error. The ρ EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS indicated that for a fixed Rater and a fixed Time point, the reliability across those sexual characteristics was .692. The ρ STABILITY and EQUIVALENCE indicated that for a fixed Rater, the reliability across time and across sexual characteristics was .690 . The ρ RATERS indicated that for a fixed Sexual Characteristic and a fixed time point, the reliability across raters was .799. The ρ RATERS value of .799 demonstrated that there was little error across raters. The ρ RATERS STABILITY indicated that for a fixed Sexual Characteristic, the reliability across raters and over time was .794. There was little error across raters and across time. The ρ RATERS EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS indicated that for a fixed time point, the reliability across raters and across sexual characteristics was .670. The ρ RATERS STABILITY OVER TIME and EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS indicated that the reliability across raters, across sexual characteristics, and over time was .660.

Although the preceding reliabilities are interesting for understanding the nature of sexual maturation in a scientific sense, they do not represent the practical situation. Researchers should use the SMIV regression equation as their estimate of sexual maturation. Therefore, the following reliabilities represent researchers using the average of the five sexual characteristics as an indicant of sexual maturation.

To adjust for using the average score of the five sexual characteristics as the indicant of sexual maturation, every $\sigma^2(e)$ and $\sigma^2(CS)$ variable was divided by the number (n) of sexual characteristics (Tables 4.18 and 4.19).

Table 4.18

Adjusted reliability formulas

Reliability	Formulas
1. ρ WITHIN	$\partial^2(S)/\partial^2(S)+\partial^2(e)(1/n)$
2. ρ STABILITY OVER TIME	$\partial^2(S)/\partial^2(S)+\partial^2(TS)+\partial^2(e)(1/n)$
3. ρ EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS	$\partial^2(S)/\partial^2(S)+\partial^2(CS)(1/n)+\partial^2(e)(1/n)$
4. ρ STABILITY and EQUIVALENCE	$\partial^2(S)/\partial^2(S)+\partial^2(TS)+\partial^2(CS)(1/n)+\partial^2(e)(1/n)$
5. ρ RATERS	$\partial^2(S)/\partial^2(S)+\partial^2(RS)+\partial^2(e)(1/n)$
6. ρ RATERS STABILITY OVER TIME	$\partial^2(S)/\partial^2(S)+\partial^2(RS)+\partial^2(RS)+\partial^2(e)(1/n)$
7. ρ RATERS EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS	$\partial^2(S)/\partial^2(S)+\partial^2(RS)+\partial^2(CS)(1/n)+\partial^2(e)(1.n)$
8. ρ RATERS STABILITY and EQUIVALENCE	$\partial^2(S)/\partial^2(S)+\partial^2(RS)+\partial^2(TS)+\partial^2(CS)(1/n)+\partial^2(e)(1/n)$

Table 4.19

The adjusted substituted reliabilities formulas, and their final values

Reliability	Substituted Formulas	Final value
1. ρ WITHIN	.8787/.8787+.1798(1/5)	.961
2. ρ STABILITY	.8787/.8787+.0038+.1798(1/5)	.957
3. ρ EQUIVALENCE	.8787/.8787+.2118(1/5)+.1798(1/5)	.918
4. ρ STABILITY and EQUIVALENCE	.8787/.8787+.0038+.2118(1/5)+.1798(1/5)	.915
5. ρ RATERS	.8787/.8787+.0417+.1798(1/5)	.919
6. ρ RATERS STABILITY	.8787/.8787+.0417+.0038+.1798(1/5)	.915
7. ρ RATERS EQUIVALENCE	.8787/.8787+.0417+.2118(1/5)+.1798(1/5)	.880
8. ρ RATERS STABILITY and EQUIVALENCE	.8787/.8787+.0417+.0038+.2118(1/5) +.1798(1/5)	.876

The adjusted ρ WITHIN indicated that for a fixed Rater, a fixed Sexual Characteristic, and a fixed Time point, the reliability of a rating was .961. The adjusted ρ STABILITY OVER TIME indicated that for a fixed Rater and a fixed Sexual Characteristic, the reliability across time was .957. Those raters virtually re-assessed the sexual characteristics without error when the analysis used the average score as the SMIV.

The adjusted ρ EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS indicated that for a fixed Rater and a fixed Time point, the reliability across

Sexual Characteristics was .918. The adjusted ρ STABILITY AND EQUIVALENCE indicated that for a fixed Rater, the reliability across Time and across Sexual Characteristics was .915. The adjusted ρ RATERS indicated that for a fixed Sexual Characteristic and a fixed Time point, the reliability across raters was .919. By using the average of the five sexual characteristics as the indicant of the underlying trait of sexual maturation, there was practically no error across raters. The adjusted ρ RATERS STABILITY indicated that for a fixed Sexual Characteristic, the reliability across raters and over time was .915. The adjusted ρ RATERS EQUIVALENCE ACROSS SEXUAL CHARACTERISTICS indicated that for a fixed point, the reliability across raters and across sexual characteristics was .880. The adjusted ρ RATERS STABILITY OVER TIME and EQUILVANCE ACROSS SEXUAL CHAR ACTERISTICS indicated that the reliability across raters, across sexual characteristics, and over time was .876.

Question III asked whether or not physical education teachers could reliably assess adolescent males' sexual maturation. The adjusted ρ RATERS reliability most directly answers that question. With a reliability estimate of .919, those physical education teachers did reliably assess the sexual maturation of the adolescent males in this sample.

Question IV. How does an estimate of skeletal age (ESA) that encompasses chronological age, height, weight, and SMIV compare with chronological age when classifying adolescent males for motor proficiency norms?

Before combining the 878 East Lansing subjects with the 700 Mason subjects for this analysis, it seemed appropriate to determine whether or not the two samples differed. Appendix L contains that discussion.

Because the review of literature demonstrated a void in the age intervals that adolescent males reached the various stages in the different sexual characteristics, Appendix M contains the proportion of this sample's 1578 boys who reach the five stages of the five sexual characteristics.

In order to compare ESA with chronological age (CA), it was necessary to determine the ESA regression equation. Appendix N contains that discussion. The ESA equation was ESA = 160.53 + 0.29(X(CA) - 159.27) + 0.33(X(H) - 63.01) + 0.23(X(W) - 108.30)+ 4.47(X(PH) - 2.95) + 4.26(X(P) - 3.02) - 0.83(X(A) - 1.86)+ 1.75(X(F) - 2.04) - 3.98(X(S) - 2.98). Table 4.20 shows the ESA and CA intervals that comprised the annual increment groupings.

Table 4.20

The month ranges that comprised the annual increment groups

Group	Month range
11.0	127 months to 138 months
12.0	139 months to 150 months
13.0	151 months to 162 months
14.0	163 months to 174 months
15.0	175 months to 186 months
16.0	187 months to 198 months
17.0	199 months to 210 months

Table 4.21 provides the ESA and CA within cell variances for the five motor proficiency tests. While the CA classification procedure had

Table 4.21

The ESA and CA within cell variances for the five motor proficiency tests

	Sit-ups	Pull-ups	Standing Long Jump	Vertical Jump	Shuttle Run
ESA	800.28	10.33	93.82	8.66	0.87
CA	763.55	9.97	94.97	8.82	0.86

a lower within cell variance for sit-ups and pull-ups, the ESA classification procedure had a lower within cell variance for standing long jump and vertical jump. And, the two classification procedures within cell variances for shuttle run were equal.

Table 4.22 provides the subjects' distribution for the ESA and CA groups. The ESA classification procedure dispersed the 1,578 subjects more evenly through the age groups than the CA classification procedure.

Table 4.22

The subjects distribution for the ESA and CA groups

	Sit-ups	Pull-ups	Standing Long Jump	Vertical Jump	Shuttle Run
ESA	800.28	10.33	93.82	8.66	0.87
CA	763.55	9.97	94.97	8.82	0.86

Table 4.23 through 4.27 provide the mean value progressions for the five motor proficiency tests in the ESA and CA groups.

Table 4.23

The mean value progressions for sit-ups in the ESA and CA groups and the between-groups increments

	Sit-ups						
	11.0	12.0	13.0	14.0	15.0	16.0	17.0
ESA	57.38	60.31	69.86	73.62	79.36	82.93	63.12
increments	(2.93)	(9.55)	(3.76)	(5.74)	(3.57)	(-19.81)	
CA	51.59	54.81	67.45	76.13	85.25	79.94	91.20
increments	(3.22)	(12.64)	(8.68)	(9.12)	(-5.31)	(11.26)	

Table 4.24

The mean value progressions for pull-ups in the ESA and CA groups and the between-groups increments

	Pull-ups						
	11.0	12.0	13.0	14.0	15.0	16.0	17.0
ESA	3.43	3.17	2.90	3.63	5.08	5.85	2.82
increments	(-0.26)	(-0.27)	(0.73)	(1.45)	(0.77)	(-3.03)	
CA	2.82	2.47	2.84	3.29	3.97	4.01	2.51
increments	(-0.35)	(0.37)	(0.45)	(0.68)	(0.04)	(-1.50)	

Table 4.25

The mean value progressions for standing long jump in the ESA and CA groups and the between-groups increments

	Standing Long Jump						
	11.0	12.0	13.0	14.0	15.0	16.0	17.0
ESA	59.68	61.39	63.76	67.32	71.46	74.43	71.18
increments	(1.71)	(2.37)	(3.56)	(4.14)	(2.97)	(-2.25)	
CA	57.44	59.86	64.52	68.19	72.21	73.13	80.60
increments	(2.42)	(4.66)	(3.67)	(4.02)	(0.92)	(7.47)	

Table 4.26

The mean value progressions for vertical jump in the ESA and CA groups and the between-groups increments

	Vertical Jump						
	11.0	12.0	13.0	14.0	15.0	16.0	17.0
ESA	12.24	12.89	13.67	14.81	15.86	16.49	15.65
increments	(0.65)	(2.28)	(1.14)	(1.05)	(0.63)	(-0.84)	
CA	11.59	12.68	13.73	15.20	15.95	15.72	15.60
increments	(1.09)	(1.05)	(1.47)	(0.75)	(-0.23)	(-0.12)	

Table 4.27

The mean value progressions for shuttle run in the ESA and CA groups and the between-groups increments

	Shuttle Run						
	11.0	12.0	13.0	14.0	15.0	16.0	17.0
ESA	11.32	11.15	10.95	10.70	10.53	10.56	11.06
increments	(-0.17)	(-0.20)	(-0.25)	(-0.17)	(0.03)	(0.50)	
CA	11.46	11.33	10.07	10.64	10.62	10.54	10.04
increments	(-0.13)	(-0.44)	(-0.25)	(-0.02)	(-0.08)	(-0.50)	

While for the 11.0, 12.0, 13.0, and 16.0 ages the ESA groups had higher sit-ups mean values, for the 14.0, 15,0, and 17.0 ages the CA groups had higher sit-up mean values. Whereas, the ESA groups increments advanced positively until the final age group, the CA group increments stopped advancing positively at the 16.0 age group.

For every age group, the ESA groups had higher pull-ups mean values. Whereas the ESA groups advanced positively only between the 13.0 and 16.0

age groups, the CA groups advanced positively between the 12.0 and 16.0 age groups.

While for the 11.0, 12.0, and 16.0 ages the ESA groups had higher standing long jump mean values, for the 13.0, 14.0, 15.0, and 17.0 ages the CA groups had higher standing long jump mean values. Whereas the ESA group increments advanced positively until the 17.0 age group, the CA group advanced positively throughout the seven age groups.

While for the 11.0, 12.0, 16.0, and 17.0 ages the ESA groups had higher vertical jump mean values, for the 13.0, 14.0, and 15.0 ages the CA groups had higher vertical jump mean values. Whereas, the ESA group increments advanced positively until the final age group, the CA group increments stopped advancing positively at the 16.0 age group.

While for the 13.0, 14.0, 16.0, and 17.0 ages the ESA groups had higher shuttle run mean values, for the 11.0, 12.0, and 15.0 ages the CA groups had higher shuttle run mean values. Relative to shuttle run scores, higher values indicated poorer scores. Whereas the ESA group increments advanced negatively until the 16.0 age group, the CA group advanced negatively throughout the seven age groups.

To answer question IV about whether or not a ESA classification procedure is superior to the CA classification procedure, the within-cell variance analysis determined no clear-cut advantage for either classification procedure. With the possible exception of the standing long jump mean values, the mean values progression analyses also failed to demonstrate a clear cut superiority for either the ESA or the CA grouping procedure. However, Crampton, Dimoch, Greulich, and other researchers have shown that while adolescent males are the same chronological age, they may be of quite different biological ages. Therefore, the fact that the ESA classification procedure dispersed the boys throughout the adolescent age period demonstrates that it might be equating the boys on the basis of biological age. Therefore, the ESA grouping might be demonstrated that the differences in the motor proficiency abilities is due to higher quality of effort or superior athletic competencies, and not biological maturation differences.

While the preceding discussion completed the four questions posed in this dissertation, the author pursued an additional classification procedure that the results suggested. Even though the analysis is far from complete, because that new procedure presented some intriguing findings, Appendix O provides the concepts upon which the investigation continues.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The purposes of this study were 1) to determine the proportion of the variance in adolescent males' skeletal age explained by increments in their chronological age, height, weight, and sexual maturation and 2) to determine how an estimate of skeletal age (ESA) compared with chronological age (CA) when classifying adolescent males for motor proficiency norms.

Increments in chronological age, height, and weight explained 45 percent, 20 percent, and 6 percent, respectively, of the variance in skeletal age. The Sexual Maturation Indicator Value estimate (SMIV) explained an additional 12 percent. In a regression equation estimate of skeletal age (ESA), those four independent variables explained 83 percent of the variance in the skeletal age of the adolescent males in this sample.

By using the variances for the cells identically determined from chronological age (CA) and ESA, neither index demonstrated a clear superiority. However, if researchers have been correct when they claim that advanced biological maturation provides adolescent males with a certainty of athletic superiority, then, because the ESA index is more likely to eliminate that advantage, ESA may classify adolescent males for motor proficiency norms better than CA.

Conclusions

The results suggest the following conclusions;

a) Without special training or competence, Boys Physical Education teachers can reliably assess the sexual maturation of adolescent males.

b) The author's SMIV checklist provides reliable ratings for the facial hair, axillary hair, pubic hair, penis, and scrotum and testes development of adolescent males.

c) Because of the high degree of intercorrelation between penis and scrotum and testes development, one of those ratings is not necessary in the SMIV regression equation.

d) Because the reliability of the penis rating exceeded the scrotum and testes rating, the SMIV regression equation that raters should use is;

$$SMIV = 126.47 + 1.74(X(F)) + 1.20(X(A)) + 2.75(X(PH)) + 6.68(X(P))$$

Where; (F) is the facial hair rating

(A) is the axillary hair rating

(PH) is the pubic hair rating, and

(P) is the penis rating.

e) Because 1) ESA classifies boys on the basis of biological maturation and 2) ESA demonstrated within-cell variances equal to or slightly better than chronological age, when classifying adolescent males for motor proficiency norms, ESA is superior to chronological age.

f) The ESA regression equation that researchers should use is;

$$ESA = (-16.15) + .277(X(CA) + .212(X(H)) + .220(X(W)) + .594(X(SMIV))$$

Where; (CA) is chronological age in months

(H) is height in inches

(W) is weight in pounds, and

(SMIV) is the Sexual Maturation Indicator Value estimate in months.

Recommendations

Based on this investigation, the author recommends;

a) Replicate the first part of this study with a sample size much larger than 30 subjects.

b) The high degree of intercorrelation between penis development and scrotum and testes development indicates that the SMIV checklist should combine those variables into a single genitalia variable.

c) As demonstrated in Appendix O, workers should continue the search for a classification index that eliminates the effect of biological maturation on the outcome of athletic competitions for adolescent males.

d) Sponsors of athletic competitions for adolescent males should re-examine their competitions in light of the effect of biological maturation and either 1) support research for a classification system that eliminates the effect of biological maturation, or 2) adjust the rules of the competition such that they eliminate the effect of biological maturation.

REFERENCES CITED

REFERENCES CITED

Acheson, Roy M. A method of assessing skeletal maturity from radiographs J ANAT 88:498-508 1954.

American Association for Health, Physical Education and Recreation AAHPER Youth Fitness Test Manual Department of National Education Association 1201 Sixteenth Street, N.W. Washington, D.C. 20036 1965.

Baldwin, Bird T. The growth of children from birth to maturity UNIV IOWA STUDIES IN CHILD WELFARE 1(1):167-187 1921.

Baldwin, Bird T. and Lorle I. Stechler Mental growth curve of normal and superior children UNIV IOWA STUDIES IN CHILD WELFARE 2(1): 1-61 1922.

Baldwin, Bird T. The determination of sex maturation in boys by a laboratory method J COMP PSY 8:39-43 1928.

Baldwin, Bird T. Laura M. Busby, and Helen V. Garside Anatomic growth of children UNIV IOWA STUDIES IN CHILD WELFARE 4(1):1-88 1928.

Bardeen, C.R. The relation of ossification to physiological development ARCH PHYS MED REHAB 2(6):1-8 1921.

Bayley, Nancy Skeletal X-rays as indicators of maturity J CONSULT PSY 4:69-73 1940.

Bayley, Nancy and Samuel R. Pinneau Tables for predicting adult height from skeletal age: revised for use with the Greulich-Pyle hand standards J PED 40:423-441 1952.

Bookwalter, Karl W. A critical evaluation of some of the existing means of classifying boys for physical education RES QUART 10(3):119-127 1939.

Bryan, A. Hughes and B.G. Greenberg Methodology in the study of physical measurements of school children HUM BIOL 24:117-144 1952.

Buehl, Clara C. and S. Idell Pyle The use of age at first appearance of three ossification centers in determining the skeletal status of children J PED 21:335-342 1942.

Carter, Thomas M. Technique and devices used in radiographic study of the wrist bones of children J ED PSY 17:237-247 1926.

Catchpole, H.R., W.W. Greulich, and R.T. Sollenberger Urinary excretion of follicle stimulating hormone in young and adolescent boys AM J PHYSIOL 123:32-33 1938.

Cattel, Psych Preliminary report on the measurement of ossification of the hand and wrist HUM BIOL 6(3):454-471 1934.

Chang, K.S. Francis, Marjorie M.M.C. Lee, and S.J. Chan Sexual maturation of Chinese boys in Hong Kong PED 34:804-811 1966.

Clarke, H. Harrison and Kay H. Peterson Contrast of maturational, structural, and strength characteristics of athletes and non-athletes 10 to 15 years of age RES QUART 32:163-176 1961.

Clarke, H. Harrison and James C.E. Harrison Differences in physical and motor traits between boys of advanced, normal, and retarded puberty RES QUART 33:13-25 1962.

Dimock, Hedley A research in adolescence CHILD DEV 6:177-195, 1935.

Dorfman, Ralph I., William Walter Greulich and Charles I Solomon The excretion of androgenic and estrogenic substances in the urine of children ENDO 21:741-743 1937.

Douglas, J.W.B. and J.M. Ross Age of puberty related to educational ability, attainment, and school leaving age J CHILD PSY PSYCHIAT 5(4):185-196 1964.

Edon, Gali Anovah: A fortran IV program to perform n-way analysis of variance Occasional Paper 27 Office of Research Consultation School for Advanced Studies College of Education Michigan State University May 20, 1976.

Ellis, Richard W.B. Height and weight in relation to onset of puberty in boys ARCH DIS CHILDH 21:181-189 1946.

Espenschade, Anna S. Motor performance in adolescence: including the study of relationships with measures of physical growth and maturity MONO SOC RES CHILD DEV 5(1):1-126 1940.

Espenschade, Anna S. Restudy of relationships between physical performances of school children and age, height, and weight RES QUART 34:144-153 1963.

Falkner, Frank Office measurement of physical growth PED CLIN N.A. 8(1):13-18 1961.

Flory, Charles D. Sex differences in skeletal development CHILD DEV 6:205-212 1935.

Flory, Charles D. Predicting puberty CHILD DEV 6:1-6 1935.

Flory, Charles, D. Osseous development in the hand as an index of skeletal development MONO SOC RES CHILD DEV 1(3):1-141 1936.

Freeman, Frank N. and Thomas M. Carter A new measure of the development of the carpal bones and its relation to the physical and mental development J ED PSYCH 15(5):257-270 1924.

Gallagher, J. Roswell and Carl C. Seltzer Maturity ratings and the prediction of height of short 14-year-old boys PED 27:984-989 1961.

Gates, Arthur I. The nature and educational significance of physical status and of mental, physiological, social, and emotional maturity J ED PSYCH 15:329-358 1924.

Greulich, William Walter, Ralph I Dorfman, Hubert R. Catchpole, Charles I. Solomon, and Charles S. Culotta Somatic and endocrine studies of puberal and adolescent boys MONO SOC RES CHILD DEV 7(3):1-84 1942.

Greulich, William W. The rational of assessing the developmental status of children from roentgenograms of the hand and wrist CHILD DEV 21:33-44 1950.

Gross, Elmer A. and Jerome A. Casciani The value of age, height, and weight as a classification device for secondary school students in the seven AAHPER Youth Fitness Tests RES QUART 33:51-58 1962.

Hellman, Milo Ossification of epiphysial cartilages in the hand AM J PHYS ANTHRO 11:223-257 1928.

Hogben, Hamilton, J.A.H. Waterhouse, and Lancelot Hogben Studies in puberty BRIT J SOC MED 2:29-42 1948.

Hunsicker, Paul and Guy G. Reiff AAHPER Youth Fitness Test Manual 1976.

Hunt, Edward E., Grace Cooke, and J. Roswell Gallagher Somatotype and sexual maturation in boys: a method of developmental analysis HUM BIOL 30:73-91 1958.

Kelly, Harriet J. Anatomic Age and its relation to stature UNIV IOWA STUDIES IN CHILD WELFARE 12(5):1-38 1937.

Krans, Hans and Rush P. Hirshland Muscular fitness and health JOPHER 24:17-19 1953.

Krans, Hans and Ruth P. Hirshland Minimum muscular fitness tests in school children RES QUART 25:178-188 1954.

Kubitscheck, Paul E. Sexual development of boys with special reference to the appearance of the secondary sexual characters and their relationship to structural and personality types J NERV MENT DIS 76:425-451 1932.

Long, Eli and E.W. Caldwell Some investigations concerning the relation between carpal ossification and physical and mental development AM J DIS CHILD 1:113-138 1911.

Lowell, Frances and Herbert Woodrow Some data on anatomical age and its relation to intelligence PED SEM 29:1-15 1922.

Marshall, Michael G. An investigation of the association between sexual maturation and physical growth and motor proficiency in males M.A. Degree Thesis Michigan State University 1967.

Marshall, W.A. and J.M. Tanner Variations in the pattern of pubertal changes in boys ARCH DIS CHILDH 45:13-23 1970.

McElhannon, Joseph C. The invalidity of the inspectional method of ranking radiographs of the carpal area of the wrists J ED PSY 17:77-85 1926.

McSweeney, Maryellen and William H. Schmidt Quantal response techniques for random predictor variables J ED STAT 2(4):257-288 1977.

Millman, Jason and Gene V. Glass Rules of thumb for writing the ANOVA table J ED MEA 4(2):41-51 1967.

Meyzi, Olcay, Hulya Alp, Ayfer Yalcindag, Suleyman Yakacikli and Alper Orphon Sexual maturation in Turkish boys ANN HUM BIOL 2(3): 251-259 1975.

Nicolson, Arline B. and Charles Hanley Indices of physiological maturity: derivation and interrelationships CHILD DEV 24(1):3-38 1953.

Prescott, Daniel Alfred The determination of anatomical age in school children and its relation to mental development HARV MONO IN EDUCATION Series 1 Number 5 1923.

Pryor, J.W. Development of the bones of the hand: as shown by the X-ray method BULL STATE COLL KENTUCKY 2(5):1-48 1905.

Pryor, J.W. Ossification of the epiphyses of the hand: X-ray method BULL STATE COLL KENTUCKY 3(4):1-35 1906.

Pryor, J.W. The hereditary nature of variation in the ossification of bones ANAT REC 1:84-88 1907.

Pryor, J.W. The chronology and order of ossification of the bones of the human carpus BULL STATE UNIV KENTUCKY 1(7):1-26 1908.

Pryor, J.W. Some observations on the ossification of the bones of the hand BULL UNIV KENTUCKY 8(11):1-67 1916.

Pryor, J.W. Differences in the time of development of centers of ossification in the male and female skeleton ANAT REC 25:257-273 1923.

Pryor, J.W. Time of ossification of the bones of the hand of the male and female and union of epiphyses with the diaphyses AM J PHYS ANTHROP 8(4):401-410 1925.

Pryor, J.W. Difference in the ossification of the male and female skeleton J ANAT 62:499-506 1928.

Pyle, S. Idell and Camille Menino Observations on estimating skeletal age from the Todd and the Flory bone Atlases CHILD DEV 10:27-34 1939.

Pyle, S. Idell, Alice M. Waterhouse, and William Walter Greulich A radiographic standard of reference for the growing hand and wrist The Press of Case Western Reserve University Year Book Medical Publishers, Inc. 35 East Wacker Drive Chicago 1971.

Rarick, Lawrence, Ionel Rapaport, and Vern Seefeldt Bone development in Down's disease AM J DIS CHILD 107:7-13 1964.

Rarick, G. Lawrence and Frank L. Smoll Stability of growth in strength and motor performance from childhood to adolescence HUM BIOL 39:295-306 1967.

Reynolds, Earle L. and Janet V. Wines Physical changes associated with adolescence in boys AMER J DIS CHILD 82:529-546 1948.

Richey, Herman G. The relation of accelerated, normal, and retarded puberty to the height and weight of school children MONO SOC RES CHILD DEV 2(1):1-67 1937.

Rotch, Thomas Morgan Chronologic and anatomic age in early life JAMA 51:1197-1205 1908.

Rotch, Thomas Morgan A study of the development of the bones in childhood by the Roentgen method, with the view of establishing a developmental index for the grading of and the protection of early life ASSN AM PHYSICIANS TRANS 24:603-624 1909.

Rotch, Thomas Morgan Roentgen ray methods applied to the grading of early life AM PHYS ED REV 15:396-421 1910.

Rotch, Thomas Morgan and Harold Wellington Smith A study of the development of the epiphyses of the hand and wrist for the purpose of classifying the cadets at Annapolis ASSN AM PHYSICIANS TRANS 25:200-211 1910.

Sawtell, Ruth Otis Ossification and growth of children from one to eight years of age AM J DIS CHILD 37:61-87 1929.

Sawtell, Ruth Otis Sex differences in the bone growth of young children AM J PHYS ANTHROP 12:293-302 1929.

Scammon, Richard E. and Gordon H. Scott The technique of determining irregular areas in morphological studies ANAT REC 35:269-277 1927.

Schiefley, Verda M. and William H. Schmidt Jeremy D. Finn's Multivariance, Univariate and Multivariate Analysis of Variance, Covariance, and Regression Occasional Paper 22 Office of Research Consultation School for Advanced Studies College of Education Michigan State University October 1973.

Schonfeld, William A. and Gilbert W. Beebe Normal growth and variation in the male genitalia from birth to maturity J UROL 48:759-777 1942.

Schonfeld, William A Primary and secondary sexual characteristics AM J DIS CHILD 65:535-549 1943.

Schonfeld, William A. Management of male pubescence JAMA 121(3):177-182 1943.

Seefeldt, Vernal D. A longitudinal study of physical growth and development in Down's syndrome unpublished Doctoral dissertation University of Wisconsin 1966.

Smith, Harold W. Rotch method of roentgenographic age determination US NAVAL MED BULL 7(1):1-20 1913.

Stevenson, Paul H. Age order of epiphyseal union in man AM J PHYS ANTHRO 7(1):53-93 1924.

Tanner, J.M. Growth at adolescence Charles C. Thomas Springfield, Illinois 1955.

APPENDICES

APPENDIX A

SEXUAL MATURATION INDICATOR VALUE (SMIV) CHECKLIST

SEXUAL MATURATION INDICATOR VALUE (SMIV) CHECKLIST

STUDENT________________________________ DATE____________________

BIRTH DATE_____________________________ SCHOOL__________________

Facial Hair (Ask the boys whether or not they shave.)

_____ No facial hair
_____ Unpigmented hair at the sideburns or the upper lip
_____ Downy, lightly pigmented hair at the sideburns or the upper lip
_____ Lightly pigmented, coarse hair at the sideburns or the upper lip
_____ Terminal sideburns or mustache

Axillary Hair

_____ No hair
_____ Slight, unpigmented hair
_____ Downy, lightly pigmented hair in one or both axilla
_____ Small area (½ to 1 inch) of pigmented, coarse hair in both axilla
_____ Coarse, curly, well-developed hair covers both axilla

Pubic Hair

_____ No hair
_____ Downy, unpigmented hair
_____ Sparse, pigmented hair, note that it is straight
_____ Small area (½ to 1 inch) of curly, coarse pigmented hair
_____ Curly, coarse, pigmented hair covers the pubic triangle

Penis Development

_____ Early childhood size and configuration
_____ Some slight enlargement evident
_____ Increase in the shaft length, glans proportionately small
_____ Proportional increase in length and diameter, but not to adult size
_____ Adult size and configuration

Scrotum and Testes Development

_____ Early childhood size and configuration
_____ Scrotum and testes visibly enlarged, one testes may show lowered
_____ One testes definitely has descended lower than the other testes
_____ Scrotal sac has darkened, but the size is still less than adult
_____ Adult size and configuration of the scrotal sac

SMIV formula: $SMIV = X + (Y - 4X)/10$

Where X is the Pubic Hair rating and Y is the sum of the Facial Hair, Axillary Hair, Penis Development, and Scrotum and Testes Development ratings.

APPENDIX B

THE PROCEDURES THAT THE RADIOLOGISTS USED TO X-RAY THE HANDS AND WRISTS OF THE SUBJECTS

The Radiology staff at Michigan State University's Olin Health Center used their general radiographic unit to X-ray both hands and wrists of the 30 boys. The radiology assistant asked the boys to place both hands on the X-ray plate, spread their fingers slightly, and gently push their palms to the surface. With the General Radiographic Unit 42 inches above the X-ray plate, the MAX set at 60 milliroentgens and, depending on the thickness of the youngster's hand, they set the milli-voltage between 66 and 70, and exposed the boys' hands and wrists to the X-ray cassette. To prevent fuzziness due to the boys' moving their hands and to minimize the X-ray exposure time, the radiologist set the exposure timer to its highest setting or less than one-half second.

APPENDIX C

THE PROCEDURES THAT PROFESSOR SEEFELDT USED TO ASSESS THE HAND-WRIST X-RAYS AND HIS TEST-RETEST RELIABILITIES

While a graduate student at the University of Wisconsin, Professor Vern Seefeldt assisted Lawrence Rarick and Ionel Rapaport in a study of Down's disease (Rarick, Rapaport, and Seefeldt, 1964). Seefeldt assessed the hand-wrist X-rays. To determine Seefeldt's ability to assess hand-wrist X-rays Rarick and Rapaport required that he compare his readings with experts. They found that Seefeldt showed no systematic bias. Professor Seefeldt's test-retest correlations were .988 for 10 year old children and .997 for 14 year old children. In his doctoral dissertation, Professor Seefeldt's test-retest correlations were .978 for 9-10 year old children and .990 for 13-14 year old children (Seefeldt, 1966). From those experiences, Professor Seefeldt became a skilled assessor of hand-wrist X-rays.

Professor Seefeldt uses the following methodology;

1) Search through the Greulich-Pyle male standards until you locate the two standards between which the X-ray fits,

2) Following the order listed in Figure C-1, assess each of the twenty-eight sites individually. Interpolate between the ages of the two standards and assign a month value to each site.

3) Average the values you assign to the twenty-eight sites for your skeletal age estimate.

Professor Seefeldt's skeletal age test-retest correlation for the 30 X-rays in this study was .96. Table C.1 provides his test-retest correlations for each of the twenty-eight sites.

Table C.1

Professor Seefeldt's test-retest correlations for the twenty-eight sites

N	Site	Correlation
1	Distal radial epiphysis	.933
2	Distal ulnar epiphysis	.921
3	Capitate	.889
4	Hamate	.895
5	Triquentral	.909
6	Lunate	.921
7	Schaphoid	.937
8	Trapezium	.931
9	Trapezoid	.942
10	First metacarpal epiphysis	.929
11	Second metacarpal epiphysis	.938
12	Third metacarpal epiphysis	.940
13	Fourth metacarpal epiphysis	.943
14	Fifth metacarpal epiphysis	.948
15	Epiphysis of the first proximal phalange	.965
16	Epiphysis of the second proximal phalange	.946
17	Epiphysis of the third proximal phalange	.947
18	Epiphysis of the fourth proximal phalange	.945
19	Epiphysis of the fifth proximal phalange	.949
20	Epiphysis of the second middle phalange	.950
21	Epiphysis of the third middle phalange	.948
22	Epiphysis of the fourth middle phalange	.956
23	Epiphysis of the fifth middle phalange	.938
24	Epiphysis of the first distal phalange	.958
25	Epiphysis of the second distal phalange	.952
26	Epiphysis of the third distal phalange	.963
27	Epiphysis of the fourth distal phalange	.964
28	Epiphysis of the fifth distal phalange	.960

APPENDIX D

THE PROCEDURES THAT THE RADIOLOGISTS USED TO MEASURE THE HEIGHT AND WEIGHT OF THE SUBJECTS

Height

Barefoot, the boys stood free on the Olin Health Center's sliding anthropometer scale while a radiology assistant lowered the horizontal bar to the midsagittal vertex of the head. Another radiology assistant recorded the height to the nearest one-quarter inch.

Weight

Naked, except for their underwear, the boys stood on the Olin Health Center's balance arm weight scale while a radiology assistant determined their weight. Another radiology assistant recorded the weight to the nearest one-half pound.

APPENDIX E

THE PROCEDURES THAT THE FOUR TEACHERS AND THE AUTHOR USED TO ASSESS THE SEXUAL MATURATION OF THE 30 SUBJECTS WHO REPRESENTED A RECTANGULAR DISTRIBUTION OF SEXUAL MATURATION

Sexual Maturation

Using the SMIV checklist, the two East Lansing, Michigan and two Mason, Michigan physical education teachers and the author assessed the sexual maturation of the 30 boys twice within a thirty minute time period. East Lansing Physical Education Director James M. Oestreich randomly ordered the boys for each assessment. All five raters had some previous experience with the SMIV checklist.

Naked, except for a towel wrapped around their waists, the boys carried ten SMIV checklists. Each rater sat in his own well-lighted assessment station. When the boys entered an assessment station, they handed one checklist to the rater and stood on a tape line approximately three feet from the rater. First, the raters asked the boys whether or not they had shaved that day. Then, the raters asked the boys to hold their towels in their right hand, raise both of their arms over their head, and turn their head to the left and then to the right. While the boys continued to look to the right, the raters completed the checklist. To assess one boy took less than fifteen seconds.

APPENDIX F

THE PROCEDURES THAT THE TEACHERS USED TO MEASURE THE HEIGHT AND WEIGHT OF THE SUBJECTS

Height

Barefoot, the East Lansing boys stood free on the physical education department's sliding anthropometer scale while the physical education teacher lowered the horizontal bar to the midsagittal vertex of their heads. Assisting the teacher, a student recorded each boy's height to the nearest one-quarter inch.

Barefoot, the Mason boys stood with their heels and backs against a wall while the physical education teacher placed a level on the mid-sagittal vertex of their heads. Assisting the teacher, a student recorded each boy's height to the nearest one-quarter inch.

Weight

Naked, except for either their underwear or athletic supporters, the East Lansing boys stood on the physical education department's balance arm weight scale while the teacher determined their weight. Assisting the teacher, a student recorded each boy's weight to the nearest one-half pound.

Naked, except for either their underwear or athletic supporter, the Mason boys stood on a calibrated bathroom scale while the teacher determined their weight. Assisting the teacher, a student recorded each boy's weight to the nearest one-half pound.

APPENDIX G

THE PROCEDURES THAT THE TEACHERS USED TO TEST THEIR STUDENTS ON THE AAHPER'S MOTOR PROFICIENCY TESTS

Sit-ups

While in a prone position on the floor, the boys extended their legs and abducted their feet to shoulder width. With a partner grasping their ankles and holding their heels against the floor, the boys placed their hands behind their neck and interlaced their fingers. Upon the teacher's command of "Go," the boys sat up, touched an elbow to the knee on the opposite side, and returned to the prone position. By alternating their elbows and the knees, the boys repeated this movement to a maximum of sit-ups. The boys could not rest between individual sit-ups. If a boy rested too long between sit-ups or stopped entirely, the teacher told the boy's partner to record the boy's sit-up score on the class sit-up sheet.

Pull-ups

The East Lansing boys used a hand-over-hand traveling ladder for their pullups. The Mason boys used a chinning bar for their pullups. With their hands pronated (the backs of their hands toward their face) and their feet off of the floor, the boys assumed a straight arm-straight leg hanging position. Upon their teacher's or another student's command of "Go," by flexing only the elbow joint, the boys elevated their body until their chin cleared the handhold bar. Each time the boys' chin cleared the bar, they received credit for one pullup. To begin another pullup, the boys had to lower their body until they fully extended their elbow joints. When the boys could no longer raise their chin above the bar, they released the handhold bar and recorded their pullup score on the class score sheet.

Standing Long Jump

The physical education teacher placed perpendicular strips of athletic tape on the gymnasium floor. A three foot strip of tape functioned as the standing long jump starting line. A perpendicular strip of tape extended ninety-six inches from the middle of the starting tape, and it functioned as the measurement tape. Beginning thirty-six inches from the starting tape, the teacher marked one inch intervals up to ninety-six inches.

Standing with their toes to the starting tape and their feet abducted to shoulder width, the boys straddled the measurement tape. By flexing their knees and simultaneously swinging their arms backward the boys began their jumping action. By simultaneously extending their knees, plantar flexing their ankles, and throwing their arms forward and upward, the boys jumped. Highly skilled boys learned to increase their distance by extending their arms just before landing. The teacher determined the distance from the starting tape to their nearest heel upon contact. They recorded the best score of three jumps.

Shuttle run

The physical education teacher placed parallel strips of athletic tape thirty feet apart on the gymnasium floor. One trip of tape acted as the Starting Tape and the other strip of tape acted as the Block Tape. Behind the Block Tape the teacher placed two 2x2x4 inch blocks of wood. With the teacher or an experienced student operating the stopwatch, the boys stood behind the starting line. Upon the timer's command of "Go," the boys sprinted to the Block Tape and grabbed one block, and sprinted back to the Starting Tape. After placing the block behind the Starting Tape, they sprinted back to the Block Tape and grabbed the second block, then, by sprinting back past the Starting Tape, they completed the shuttle run test. The timer determined their time to the nearest one-tenth of a second. They recorded the best score of three runs.

Vertical Jump

Beginning at a height of five feet off the floor, the physical education teacher placed a six foot strip of tape vertically on the gymnasium wall with markings one inch apart. To determine their reaching height, the boys stood flatfooted with one side of their body against the wall. With the hand closest to the wall they reached as high as they could. The number to which they reached became their reaching height number. To determine their jumping height, the boys moistened the fingertips of the same hand, then by pliometrically extending their knees and swinging their arms backward the boys began their jumping action. By simultaneously throwing their arms upward, extending their knees, and plantar flexing their ankles, the boys jumped as high as they could. At the peak of their jump, the boys touched the measurement tape with the moistened fingertips. The height where the fingertips touched the tape became the boys' jumping height. By subtracting reaching height from jumping height, the teacher or an experienced student assistant determined the vertical jump score. They recorded the best score of three jumps.

APPENDIX H

THE PROCEDURES THAT THE TEACHERS USED TO ASSESS THE SEXUAL MATURATION OF THEIR SUBJECTS

Using the SMIV checklist, physical education teachers from East Lansing and Mason Junior High Schools assessed five primary and secondary sex characteristics. In conjunction with their AAHPER Youth Fitness testing week, the physical education teachers daily asked seven or eight boys to report to them after physical education class. In East Lansing, the physical education office served as the station for the assessment of sexual maturation. In Mason, the towel disposal room beside the showers served as the station for the assessment of sexual maturation. Both locations were well lighted.

Naked, except for a towel wrapped around them, the boys entered the station and stood on a tape line approximately three feet from the teachers. After writing the boy's name on the SMIV checklist, the teachers asked the boys whether or not they shaved that day. Then, the teachers asked the boys to hold the towel in their right hand, to raise both arms over their head, and to turn their head to the left and then to the right. While the boys continued to look to the right, the teachers completed the checklist. To assess one boy and explain his SMIV score required less than thirty seconds.

APPENDIX I

THE STEP-WISE REGRESSIONS FOR THE REORDERED CHRONOLOGICAL AGE, HEIGHT, AND WEIGHT INDEPENDENT VARIABLES

Table I-1

After height, the percent of additional variance accounted for by chronological age and the significance level

	Percent of Additional Variance	Significance Level
Height	61.931	
Chronological Age	2.856	P less than .1506

Table I-2

After chronological age, the percent of additional variance accounted for by weight and the significance level

	Percent of Additional Variance	Significance Level
Chronological Age	44.937	
Weight	23.191	P less than .0002

Table I-3

After weight, the percent of additional variance accounted for by chronological age and the significance level

	Percent of Additional Variance	Significance Level
Weight	49.559	
Chronological Age	18.569	P less than .0005

Table I-4

After height, the percent of additional variance accounted for by weight and the significance level

	Percent of Additional Variance	Significance Level
Height	61.931	
Weight	3.978	P less than .0872

Table I-5

After weight, the percent of additional variance accounted for by height and the significance level

	Percent of Additional Variance	Significance Level
Weight	49.559	
Height	16.350	P less than .0013

Table I-6

After chronological age and weight, the percent of additional variance accounted for by height and the significance level

	Percent of Additional Variance	Significance Level
Chronological Age	44.937	
Weight	23.191	P less than .0002
Height	2.625	P less than .1387

After accounting for their chronological age contribution, the boys' height did have a significant additional contribution unique to it. However, after accounting for the contribution of height, chronological age did not have a significant additional contribution unique to it. The reason for this is that chronological age is very highly correlated with height, such that, after accounting for the common contribution of height and chronological age, chronological age has nothing left to add.

Regardless of the order of chronological age and weight, both variables had a significant contribution unique to them.

After accounting for height, weight did not have a significant addition. The high intercorrelation between height and weight also caused this finding. However, after accounting for weight, height did have a significant additional contribution. Height has a contribution that is unique to it beyond its common contribution with weight.

APPENDIX J

THE PROCEDURES FOR DETERMINING A SEXUAL MATURATION INDICATOR VALUE(SMIV) REGRESSION EQUATION

Before determining a SMIV regression equation, we should examine the sample correlation matrix for the five sexual characteristics (Table J-1).

Table J-1

The correlation matrix for the five sexual characteristics

	Facial Hair	Axillary Hair	Pubic Hair	Penis Dev.	Scrotum Dev.
Facial Hair	1.000				
Axillary Hair	.849	1.000			
Pubic Hair	.697	.670	1.000		
Penis Dev.	.710	.671	.931	1.000	
Scrotum Dev.	.787	.731	.902	.954	1.000

Ideally, in a regression equation, the independent variables should not be highly intercorrelated. However, because those variables relate to some particular variable, in this case sexual maturation, they frequently are highly intercorrelated. When two independent variables in a regression equation are highly intercorrelated, sometimes the contribution of the variable listed first devours all of the contribution of the variable listed second. In this correlation matrix, the high correlation of penis development and scrotum and testes development (.954) and pubic hair with penis development (.931) and scrotum and testes development (.902) might effect the SMIV regression equation.

To obtain a regression equation for estimating SMIV, the Multivariate Regression Analysis estimated the dependent variable (skeletal age) from the independent variables (the five sexual characteristics). Table J-2 contains the raw regression coefficients and mean values that the Multivariate Regression Analysis determined for the SMIV regression equation.

The mean deviated SMIV regression equation is:

$$SMIV = \hat{\mu}(SA) + \hat{\beta}(F)(X(F) - \hat{\mu}(F)) + \hat{\beta}(A)(X(A) - \hat{\mu}(A)) + \hat{\beta}(PH)(X(PH) - \hat{\mu}(PH)) + \hat{\beta}(P)(X(P) - \hat{\mu}(P)) + \hat{\beta}(S)(X(S) - \hat{\mu}(S))$$

Table J-2

The raw regression coefficients and mean values that the Multivariate Regression Analysis determined for the SMIV regression equation

	Raw Regression Coefficients ($\hat{\beta}$)	Mean Values($\hat{\mu}$)
Facial Hair (F)	1.94	2.04
Axillary Hair (A)	1.21	1.86
Pubic Hair (PH)	2.77	2.95
Penis Dev. (P)	7.27	3.02
Scrotum Dev. (S)	-1.06	2.98
Skeletal Age (SA)	----	160.53

The mean deviated SMIV regression equation is:

$$SMIV = \hat{\mu}(SA) + \hat{\beta}(F)(X(F)-\hat{\mu}(F) + \hat{\beta}(A)(X(A))-\hat{\mu}(A) + \hat{\beta}(PH)(X(PH)-\hat{\mu}(PH)) + \hat{\beta}(P)(X(P)-\hat{\mu}(P)) + \hat{\beta}(S(X(S)-\hat{\mu}(S))$$

To convert this mean deviated SMIV regression equation to non-mean deviated, it is necessary to solve the following SMIV $\beta(0)$ equation;

$$SMIV\ \beta(0) = \hat{\mu}(SA) - \hat{\beta}(F)\hat{\mu}(F) - \hat{\beta}(A)\hat{\mu}(A) - \hat{\beta}(PH)\hat{\mu}(PH) - \hat{\beta}(P)\hat{\mu}(P) - \hat{\beta}(S)\hat{\mu}(S)$$

After substituting the β and μ values from Table J-2 and solving the equation, SMIV $\beta(0)$ equalled 127.35.

To demonstrate the effects of increasing the ratings of the five sexual characteristics individually and to determine the common intervals and range of the SMIV regression equation, Table J-3 substitutes the ratings 1 through 5 for the sexual characteristics values. While a one-level increase in facial hair, axillary hair, pubic hair, and penis development represented monthly increases of 1.94, 1.21, 2.77, and 7.27, respectively, a one-level increase in scrotum and testes development represented a loss of 1.06 in the SMIV regression equation.

To have a decrease in SMIV when scrotum and testes development increased demonstrated a problem in the regression equation. Why did

Table J-3

Substituting selected ratings for the five sexual characteristics values into the SMIV regression equation

β(F)	(X(F))+	β(A)	(X(A))+	β(PH)	(X(PH)+	β(P)	(X(P))+	β(X)	(X(S))+	β (O) = SMIV
1.94	(1) +	1.21	(1) +	2.77	(1)	7.27	(1)	-1.06	(1) +	127.35=139.48
1.94	(2) +	1,21	(1) +	2.77	(1)	7.27	(1)	-1.06	(1) +	127.35=141.42
1.94	(1) +	1.21	(2) +	2.77	(1)	7.27	(1)	-1.06	(1) +	127.35=140.69
1.94	(1) +	1.21	(1) +	2.77	(2)	7.27	(1)	-1.06	(1) +	127.35=142.24
1.94	(1) +	1.21	(1) +	2.77	(1)	7.27	(2)	-1.06	(1) +	127.35=146.75
1.94	(1) +	1.21	(1) +	2.77	(1)	7.27	(1)	-1.06	(2) +	127.35=138.42
1.94	(2) +	1.21	(2) +	2.77	(2)	7.27	(2)	-1.06	(2) +	127.35=151.61
1.94	(3) +	1.21	(3) +	2.77	(3)	7.27	(3)	-1.06	(3) +	127.35=163.74
1.94	(4) +	1.21	(4) +	2.77	(4)	7.27	(4)	-1.06	(4) +	127.35=175.87
1.94	(5) +	1.21	(5) +	2.77	(5)	7.27	(5)	-1.06	(5) +	127.35=188.00

the beta weight for scrotum and testes development have a negative value? High intercorrelation between independent variables can cause a negative beta weight for one of the variables. Statisticians refer to this phenomenon as multi-colinearity. Multi-colinearity prevents the regression analysis from determining the singular effects of the two highly correlated independent variables.

Recall that the variables of scrotum and testes development and penis development were highly correlated (.954). The procedure that confirms the multi-colinearity problem is to eliminate one and then the other of the variables from the list of independent variables and examine the new beta weights. If the beta weights in those two lists are positive, then the intercorrelation did cause the multi-colinearity problem.

Table J-4 contains the raw regression coefficients for the two lists of independent variables. Because all of the raw regression coefficients are positive, the high intercorrelation between the penis development variable and the scrotum and testes development definitely caused the negative beta weight for the scrotum and testes development variable.

Table J-4

Raw regression coefficients for the regression analyses minus the penis development variable and minus the scrotum and testes development variable

	Minus Penis Dev. Raw Regression Coefficients	Minus Scrotum Dev. Raw Regression Coefficients
Facial Hair	1.74	0.88
Axillary Hair	1.20	1.26
Pubic Hair	2.75	5.46
Penis Dev.	6.68	----
Scrotum Dev.	----	6.05

The solution for multi-colinearity is to eliminate one of the variables. Because the two variables were highly intercorrelated, they contributed the same manifestation to the regression equation and to eliminate one of the variables does not weaken the regression equation. Therefore, the question became whether to eliminate the penis development variable or the scrotum and testes development variable.

The mean-deviated minus scrotum and testes development (MSD) regression equation is;

$$(MSD)\ SMIV = \hat{\mu}(SA) + \hat{\beta}(F)(X(F)-\hat{\mu}(F)) + \hat{\beta}(A)(X(A)-\hat{\mu}(A)) + \hat{\beta}(PH)(X()(PH)- \hat{\mu}(PH)) + \hat{\beta}(P)(X(P)-\hat{\mu}(P))$$

To convert this mean deviated (MSD) SMIV regression equation to a non-mean deviated, it is necessary to solve the following (MSD) SMIV $\beta(0)$ equation;

$$(MSD)\ SMIV\ \beta(0) = \hat{\mu}(SA) - \hat{\beta}(F)\hat{\mu}(F) - \hat{\beta}(A)\hat{\mu}(A) - \hat{\beta}(PH)\hat{\mu}(PH) - \hat{\beta}(P)\hat{\mu}(P)$$

After substituting the Table J-4 $\hat{\beta}$'s and the Table J-2 $\hat{\mu}$'s and solving the equation, (MSD) SMIV $\beta(0)$ equalled 126.47.

To demonstrate the effects of increasing the ratings of those four sexual characteristics individually and to determine the common intervals and range of the (MSD) SMIV regression equation, Table J-5 substitutes the ratings 1 through 5 for the sexual characteristics.

Table J-5

Substituting selected ratings for the four sexual characteristics values into the (MSD) SMIV regression equation.

(F)	(X(F)) +	(A)	(X(A)) +	(PH)	(X(PH)) +	(P)	(X(P)) +	$\beta(0)$	=(MSD) SMIV
1.74	(1)	1.20	(1)	2.75	(1)	6.68	(1)	126.47	= 138.84
1.74	(2)	1.20	(1)	2.75	(1)	6.68	(1)	126.47	= 140.58
1.74	(1)	1.20	(2)	2.75	(1)	6.68	(1)	126.47	= 140.04
1.74	(1)	1.20	(1)	2.75	(1)	6.68	(1)	126.47	= 141.59
1.74	(1)	1.20	(1)	2.75	(1)	6.68	(1)	126.47	= 145.42
1.74	(2)	1.20	(2)	2.75	(1)	6.68	(1)	126.47	= 151.21
1.74	(3)	1.20	(3)	2.75	(1)	6.68	(1)	126.47	= 163.58
1.74	(4)	1.20	(4)	2.75	(1)	6.68	(1)	126.47	= 175.96
1.74	(5)	1.20	(5)	2.75	(1)	6.68	(1)	126.47	= 188.32

A one level increase in facial hair, axillary hair, pubic hair, and penis development represented monthly increases of 1.74, 1.20, 2.75, and 6.68, respectively in the (MSD) SMIV.

The mean deviated minus penis development (MPD) SMIV regression equation is;

$$\text{(MPD) SMIV} = \hat{\mu}(SA) + \hat{\beta}(F)(X(F)-\hat{\mu}(F)) + \hat{\beta}(A)(X(A))-\hat{\mu}(A)) + \hat{\beta}(PH)(X(PH)-\hat{\mu}(PH)) + \hat{\beta}(S)(X(S)\hat{\mu}(S))$$

To convert this mean deviated (MPD) SMIV regression equation to a non-mean deviated, it is necessary to solve the following (MPD) SMIV $\beta(0)$ equation:

$$\text{(MPD) SMIV } \beta(0) = \hat{\mu}(SA) - \hat{\beta}(F)\hat{\mu}(F) - \hat{\beta}(A)\hat{\mu}(A) - \hat{\beta}(PH)\hat{\mu}(PH) - \hat{\beta}(S)\ \hat{\mu}(S)$$

By substituting the Table J-4 (MPD) $\hat{\beta}$'s and the Table J-2 $\hat{\mu}$'s and solving the equation, the (MPD) SMIV $\beta(0)$ equalled 122.24.

To demonstrate the effects of increasing the ratings of those four sexual characteristics individually and to determine the common intervals and range of the (MPD) SMIV regression equation, Table J-6 substitutes the ratings 1 through 5 for the sexual characteristics.

A 1 level increase in facial hair, axillary hair, pubic hair, and scrotum and testes development represented monthly increases of 0.88,

Table J-6

Substituting selected ratings for the four sexual characteristics values into the (MPD) SMIV regression equation

$\hat{\beta}$(F)	(X(F))	+ β(A)	(X(A))	+ $\hat{\beta}$(PH)	(X(PH)	+ $\hat{\beta}$(S)	(X(S))	+ $\hat{\beta}$(0)	= (MPD)SMIV
0.88	(1)	+ 1.26	(1)	+ 5.46	(1)	+ 6.05	(1)	+ 122.24	= 135.89
0.88	(2)	+ 1.26	(1)	+ 5.46	(1)	+ 6.05	(1)	+ 122.24	= 136.77
0.88	(1)	+ 1.26	(2)	+ 5.46	(1)	+ 6.05	(1)	+ 122.24	= 137.15
0.88	(1)	+ 1.26	(1)	+ 5.46	(2)	+ 6.05	(1)	+ 122.24	= 141.35
0.88	(1)	+ 1.26	(1)	+ 5.46	(1)	+ 6.05	(2)	+ 122.24	= 141.94
0.88	(2)	+ 1.26	(2)	+ 5.46	(2)	+ 6.05	(2)	+ 122.24	= 149.54
0.88	(3)	+ 1.26	(3)	+ 5.46	(3)	+ 6.05	(3)	+ 122.24	= 163.19
0.88	(4)	+ 1.26	(4)	+ 5.46	(4)	+ 6.05	(4)	+ 122.24	= 176.84
0.88	(5)	+ 1.26	(5)	+ 5.46	(5)	+ 6.05	(5)	+ 122.24	= 190.49

1.26, 5.46, and 6.05, respectively in the (MPD) SMIV.

Table J-7 contains the SMIV, (MDS) SMIV, and (MPD) SMIV values.

Table J-7

Comparing the SMIV, (MSD) SMIV, and (MPD) SMIV values

Selected Values							
(X(F))	(X(A))	(X(PH))	(X(P))	(X(S))	SMIV	(MSD)SMIV	(MPD)SMIV
(1)	(1)	(1)	(1)	(1)	139.48	138.84	135.89
(2)	(1)	(1)	(1)	(1)	141.42	140.58	136.77
(1)	(2)	(1)	(1)	(1)	140.69	140.04	137.15
(1)	(1)	(2)	(1)	(1)	142.24	141.59	141.35
(1)	(1)	(1)	(2)	(1)	146.75	145.42	----
(1)	(1)	(1)	(1)	(2)	138.42	---0	141.95
(2)	(2)	(2)	(2)	(2)	151.61	151.21	149.54
(3)	(3)	(3)	(3)	(3)	163.74	163.58	163.19
(4)	(4)	(4)	(4)	(4)	175.87	175.95	176.84
(5)	(5)	(5)	(5)	(5)	188.00	188.32	190.94

Comparing the (MSD) SMIV and (MPD) SMIV values failed to demonstrate a clear superiority for either regression equation. Therefore, the rater reliability for measuring the penis development and scrotum and testes variables will determine whether to use (MSD)SMIV or (MPD)SMIV. Nevertheless, because the original SMIV regression equation contained a negative beta weight for the scrotum and testes development did not prevent its use when analyzing the additional contribution unique to SMIV.

APPENDIX K

THE STEP-WISE REGRESSIONS FOR THE REORDERED CHRONOLOGICAL AGE, HEIGHT WEIGHT, AND SMIV INDEPENDENT VARIABLES

Table K-1

The step-wise regression for chronological age and SMIV

	Percent of Additional Variance	Significance Level
Chronological Age	44.937	
SMIV	30.031	P less than .0001

Table K-2

The step-wise regression for SMIV and chronological age

	Percent of Additional Variance	Significance Level
SMIV	68.811	
Chronological Age	6.158	P less than .0158

Table K-3

The step-wise regression for height and SMIV

	Percent of Additional Variance	Significance Level
Height	61.934	
SMIV	15.245	P less than .0003

Table K-4

The step-wise regression for SMIV and height

	Percent of Additional Variance	Significance Level
SMIV	68.881	
Height	8.377	P less than .0040

Table K-5

The step-wise regression for weight and SMIV

	Percent of Additional Variance	Significance Level
Weight	49.559	
SMIV	28.649	P less than .0001

Table K-6

The step-wise regression for SMIV and weight

	Percent of Additional Variance	Significance Level
SMIV	68.811	
Weight	9.397	P less than .0021

Regardless of the order of chronological age and SMIV, both contributed a significant additional proportion of the variance. Regardless of the order of height and SMIV, both contributed a significant additional proportion of the variance. And, regardless of the order of weight and SMIV, both contributed a significant additional proportion of the variance. The proportion of the variance in skeletal age when chronological age and SMIV were the independent variables, when height and SMIV were the independent variables, and when weight and SMIV were the independent variables was 75, 77, and 78 percent, respectively.

APPENDIX L

THE STATISTICAL ANALYSIS TO DETERMINE WHETHER OR NOT THE 878 EAST LANSING SUBJECTS DIFFERED FROM THE 700 MASON SUBJECTS

The Multivariate Analysis of Variance computer program analyzed whether or not the two school's data were significantly different from one another (Scheifley and Schmidt, 1973). In the Multivariate Analysis of Variance analysis, the F-ratio for the multivariate test of equality of the mean vectors indicates whether or not the grouping factors were significantly different. Because this analysis is very sensitive towards general F-ratio significance, the univariate F-ratios indicates the variables that caused the general F-ratio significance.

The factors upon which to group those data for the comparisons were; two levels of a SCHOOL factor (East Lansing and Mason), four levels of a GRADE factor (seventh Grade fall(7th F), seventh grade spring(7th S), eighth grade fall(8th F), and eighth grade spring(8th Sp)), and three levels of a YEAR factor (year one(1), year two(2), and year three(3)). Because East Lansing had sixth grade data and Mason had none, and Mason had ninth grade data and East Lansing had none, the analysis compared only the seventh and eighth grade subjects from both schools.

Because all of the data were cross-sectional, it was necessary to determine whether or not the boys in the six test periods differed from each other. Therefore, the first two comparisons contrasted the YEAR factor with the GRADE factor for the East Lansing boys and the Mason boys separately. Then, the third comparison contrasted the SCHOOL factor with the GRADE factor across the YEAR factor, or a triple interaction analysis.

<u>Contrast #1</u> (GRADE by YEAR for East Lansing only)

In a multi-factor design, to have the main effects and interactions orthogonal to each other, the cell sizes must be equal or proportional. Therefore, after determining the mean values for the cells containing the total sample, forty subjects from each cell were randomly selected as the analysis sample. Then, the mean values in those analysis cells were determined (Tables L-1 through L-14). The mean values in the analysis sample represent the mean values for the total sample.

Table L-1

East Lansing data's number of subjects in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	48 - 40	68 - 40	76 - 40	156 - 40
2	41 - 40	49 - 40	48 - 40	57 - 40
3	74 - 40	96 - 40	40 - 40	42 - 40

Table L-2

East Lansing data's chronological age mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	149.69-150.01	155.79-155.61	162.12-162.96	166.83-167.19
2	149.45-149.41	155.76-155.74	163.15-163.08	168.27-168.42
3	150.40-149.71	155.33-155.56	163.16-163.16	167.83-167.84

Table L-3

East Lansing data's height mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	61.27-61.18	16.35-61.53	63.69-63.93	64.54-64.79
2	59.45-59.49	61.64-62.10	62.84-62.64	64.23-64.74
3	60.17-60.44	62.73-63.27	63.15-63.15	65.16-65.17

Table L-4

East Lansing data's **height** mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	97.80-97.12	108.40-109.85	114.11-118.82	122.95-120.43
2	95.68-95.65	105.98-106.76	111.50-110.14	127.07-127.41
3	99.15-101.45	110.54-111.20	111.39-111.39	123.09-123.40

Table L-5

East Lansing data's sit-ups mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	68.87-69.15	76.31-74.85	80.09-78.93	83.82-84.23
2	43.61-42.23	74.53-75.63	69.79-60.50	78.24-77.38
3	56.97-59.05	65.17-67.73	59.97-59.98	68.72-68.33

Table L-6

East Lansing data's pull-ups mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	3.42-3.18	2.98-2.20	3.55-3.50	3.46-3.88
2	2.41-2.30	3.00-2.93	3.33-3.30	2.68-2.53
3	3.01-2.87	2.79-2.98	3.63-3.63	4.05-4.05

Table L-7

East Lansing data's standing long jump mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th spring	8th Fall	8th Spring
1	64.60-65.35	66.91-65.75	67.31-66.50	70.58-70.45
2	60.39-60.15	67.63-67.53	64.73-65.10	70.98-71.60
3	60.84-61.35	67.81-67.70	64.30-64.30	72.79-72.88

Table L-8

East Lansing data's vertical jump mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	14.79-14.73	14.50-13.93	15.66-15.58	16.76-16.58
2	13.20-13.13	14.80-14.85	14.19-14.13	15.82-16.05
3	14.04-14.05	13.69-13.73	14.06-14.60	15.93-16.13

Table L-9

East Lansing data's shuttle run mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	10.93-10.94	10.30-10.41	10.52-10.44	10.14-10.16
2	12.10-12.13	10.65-10.62	11.14-11.19	10.12-10.22
3	10.76-10.75	10.62-10.67	10.84-10.84	10.25-10.24

Table L-10

East Lansing data's facial hair mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	1.31-1.30	1.66-1.55	2.45-2.55	2.25-2.13
2	1.83-1.80	1.88-1.90	2.12-2.08	2.68-2.58
3	1.22-1.10	1.97-2.10	1.90-1.90	2.31-2.35

Table L-11

East Lansing data's axillary hair mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th spring	8th Fall	8th Spring
1	1.08-1.08	1.74-1.63	2.29-2.40	2.55-2.43
2	1.24-1.25	1.63-1.68	1.94-1.88	2.74-2.65
3	1.27-1.33	1.86-1.85	2.42-2.43	2.67-2.68

Table L-12

East Lansing data's pubic hair mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	1.83-1.88	2.84-2.90	3.17-3.30	3.38-3.45
2	1.88-1.88	2.53-2.63	3.29-3.28	3.68-3.70
3	2.34-2.43	2.92-3.00	3.23-3.23	3.76-3.78

Table L-13

East Lansing data's penis development mean value in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	2.08-2.08	2.79-2.80	3.04-3.03	3.13-3.20
2	2.17-2.15	2.51-2.63	3.02-3.03	3.58-3.63
3	2.57-2.50	2.91-3.00	3.43-3.43	3.81-3.80

Table L-14

East Lansing data's scrotum and testes development mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	2.25-2.28	2.85-2.83	3.16-3.20	3.26-3.23
2	2.10-2.10	2.69-2.75	3.08-3.05	3.67-3.70
3	2.12-2.08	2.54-2.48	2.78-2.78	3.74-3.73

Table L-15 provides the univariate F values and their associated significance levels for the East Lansing GRADE by YEAR interaction analysis.

The general F-ratio for the East Lansing GRADE by YEAR interaction was 3.0913. With a significance level less than .05, the analysis indicated that East Lansing's four levels of the GRADE factor and three levels of the YEAR factor interacted significantly. However, because the four levels of the GRADE factor contained boys aged six months apart, the GRADE by YEAR interaction was probable. Therefore, the univariate F's and their

Table L-15

The univariate F values and their associated significance levels for the East Lansing GRADE by YEAR interaction

Variables	Univariate F	P less than
Chronol. Age	.169	.9850
Height	1.419	.2057
Weight	1.132	.3426
Sit-ups	2.957	.0077
Pull-ups	1.267	.2714
Stand Lg Jp	1.764	.1048
Vertical Jp	1.879	.0827
Shuttle Run	7.793	.0001
Facial hair	5.310	.0001
Axillary hair	.871	.5158
Pubic hair	1.139	.3385
Penis dev.	.664	.6792
Scrotum dev.	2.208	.0412

significance levels became the interesting statistics.

Because the univariate F analysis contained thirteen variables, the probability level for each variable became .05/13 or .004. Figures L.1 and L.2 graphically illustrate the interactions in shuttle run and facial hair.

The second year test period 7th F boys having the very low mean score, and all three year test period 7th Sp boys having higher scores than the 8th F boys caused the GRADE by YEAR significant differences in shuttle run. The third year test period 7th SP boys having a higher score than the 8th F boys, and the first year test period F boys having a higher score than the 8th Sp boys caused the GRADE by YEAR significant differences in facial hair.

Contrast #2 (GRADE by YEAR for Mason only)

After determining the mean values for the cells containing the total sample, twenty-five subjects from each cell were randomly selected

Figure L-1

The East Lansing data's shuttle run mean values GRADE by YEAR interaction

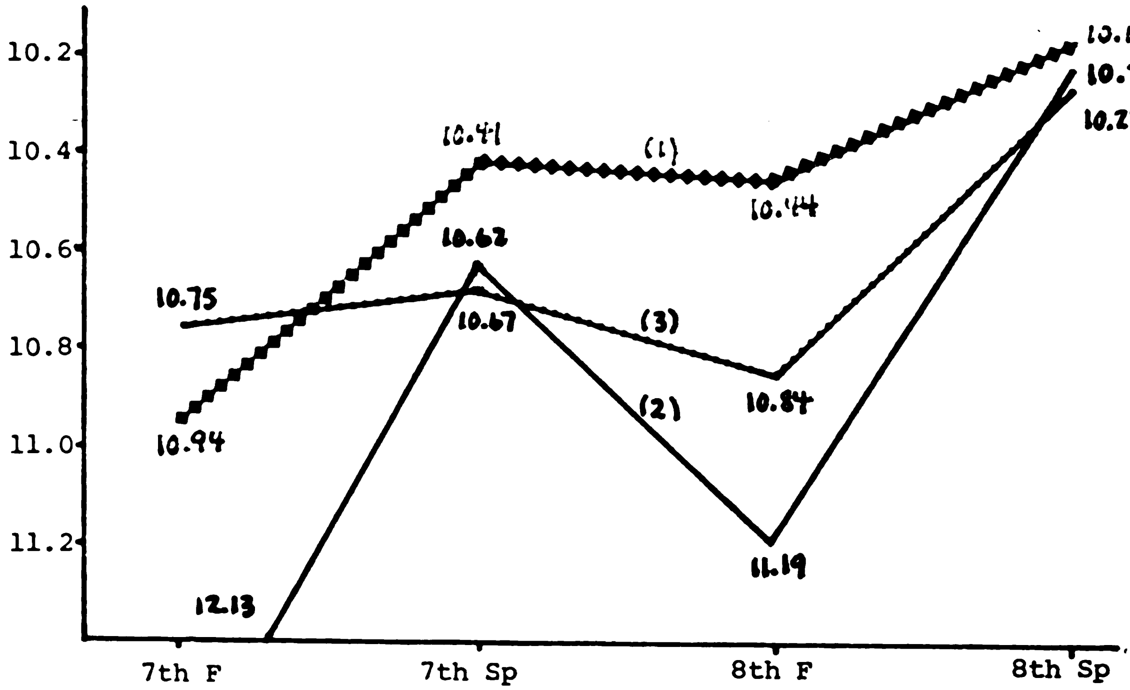
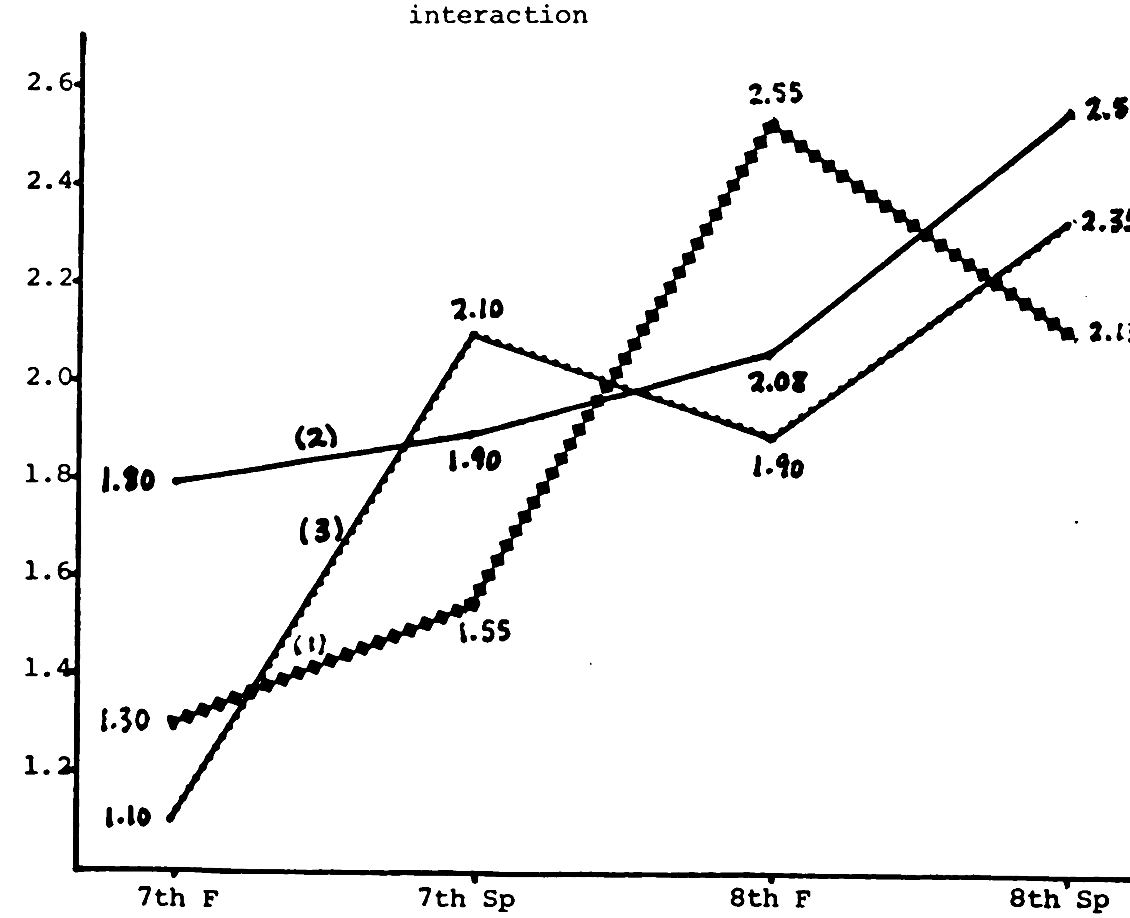


Figure L-2

The East Lansing data's facial hair mean values GRADE by YEAR interaction



as the analysis sample. Then, the mean values for those analysis cells were determined (Tables L-16 through 29). Because the analysis sample in the East Lansing cells contained forty subjects, by putting twenty-five subjects in each Mason cell, the comparison of the East Lansing and Mason data had proportional cells. Because the author randomly selected the analysis cell samples, again, the mean values in the analysis sample represent the mean values in the total sample.

Table L-16

Mason data's number of subjects in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	31-25	29-25	30-25	26-25
2	28-25	31-25	36-25	32-25
3	60-25	28-25	51-25	30-25

Table L-17

Mason data's chronological age mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	150.19-150.14	156.07-156.36	165.18-164.68	172.60-173.02
2	150.09-150.66	156.64-157.50	162.53-161.22	167.77-167.12
3	153.42-152.86	154.86-154.90	162.17-163.04	168.10-168.04

Table L-18

Mason data's height mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	60.21-60.10	62.45-62.19	63.46-63.13	65.30-65.26
2	60.42-60.48	62.28-62.42	62.94-62.91	64.07-63.95
3	60.53-61.05	61.82-61.47	62.80-62.86	64.95-64.50

Table L-19

Mason data's weight mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	95.40-95.32	109-31-110-12	126.85-123.84	123.35-123.56
2	101.07-100.84	116.95-119.72	110.62-113-34	116.61-114.22
3	103.30-108.43	108.71-105.12	116.39-112.48	126.50-125.76

Table L-20

Mason data's sit-ups mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	42.23-41.96	64.19-63.16	72.20-68.88	87.23-89.12
2	35.96-33.48	49.84-49.36	45.53-40.60	78.78-86.04
3	66.27-65.40	71.86-71.72	69.20-73.80	76.57-75.60

Table L-21

Mason data's pull-ups mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	2.77-2.88	3.93-3.84	3.40-3.00	6.35-6.36
2	3.75-3.60	2.68-2.36	3.28-3.00	5.22-5.44
3	2.93-2.28	3.04-3.32	2.76-2.44	4.17-4.54

Table L-22

Mason data's standing long jump mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th spring	8th Fall	8th Spring
1	54.35-54.64	57.48-57-48	61.90-61.84	69.81-70.00
2	59.04-58.92	60.32-59-72	60-35-59.48	68.91-68.88
3	60.53-60.56	60.14-60.60	63.57-62.92	66.50-65.28

Table L-23

Mason data's vertical jump mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	11.29-11.04	11.86-12.00	11.63-11.44	14.46-14.60
2	11.32-11.32	11.77-11.60	12.53-12.40	13.88-13.96
3	12.38-12.12	12.57-12.68	13.65-13.60	14.03-13.72

Table L-24

Mason data's shuttle run mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	11.60-11.61	12.05-12.01	11.23-11.30	11.23-11.23
2	12.05-12.07	11.56-11.58	11.54-11.59	11.24-11.12
3	11.22-11.26	11.17-11.11	10.80-10.94	10.99-11.08

Table L-25

Mason data's facial hair mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	1.39-1.40	2.34-2.32	2.67-2.52	3.11-3.20
2	1.50-1.56	1.94-1.96	2.11-1.96	2.22-2.20
3	1.42-1.32	1.71-1.72	1.47-1.56	2.23-2.20

Table L-26

Mason data's axillary hair mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th spring	8th Fall	8th Spring
1	1.52-1.40	2.17-2.24	2.93-2.84	2.85-2.92
2	1.53-1.60	2.13-2.16	1.97-2.00	2.16-2.12
3	1.67-1.56	1.93-1.88	1.88-1.84	2.33-2.40

Table L-27

Mason data's pubic hair mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	1.84-1.76	2.79-2.84	3.30-3.16	3.31-3.32
2	2.11-2.20	2.81-2.96	2.83-2.84	3.25-3.28
3	2.03-1.80	2.50-2.48	2.75-2.60	3.40-3.36

Table L-28

Mason data's penis development mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	2.26-2.24	2.41-2.40	3.33-3.20	3.23-3.24
2	2.14-2.20	2.61-2.68	2.56-2.44	3.28-3.24
3	2.00-1.80	2.25-2.24	2.71-2.68	3.10-3.00

Table L-29

Mason data's scrotum and testes development mean values in the total sample and in the analysis sample

YEAR	GRADE			
	7th Fall	7th Spring	8th Fall	8th Spring
1	2.03.2.00	2.69-2.76	3.30-3.16	3.19-3.24
2	2.18-2.28	2.45-2.52	2.67-2.64	3.06-3.04
3	1.95-1.80	2.21-2.20	2.65-2.60	3.07-2.96

Table L-30

The univariate F values and their associated significance levels for the Mason GRADE by YEAR interaction

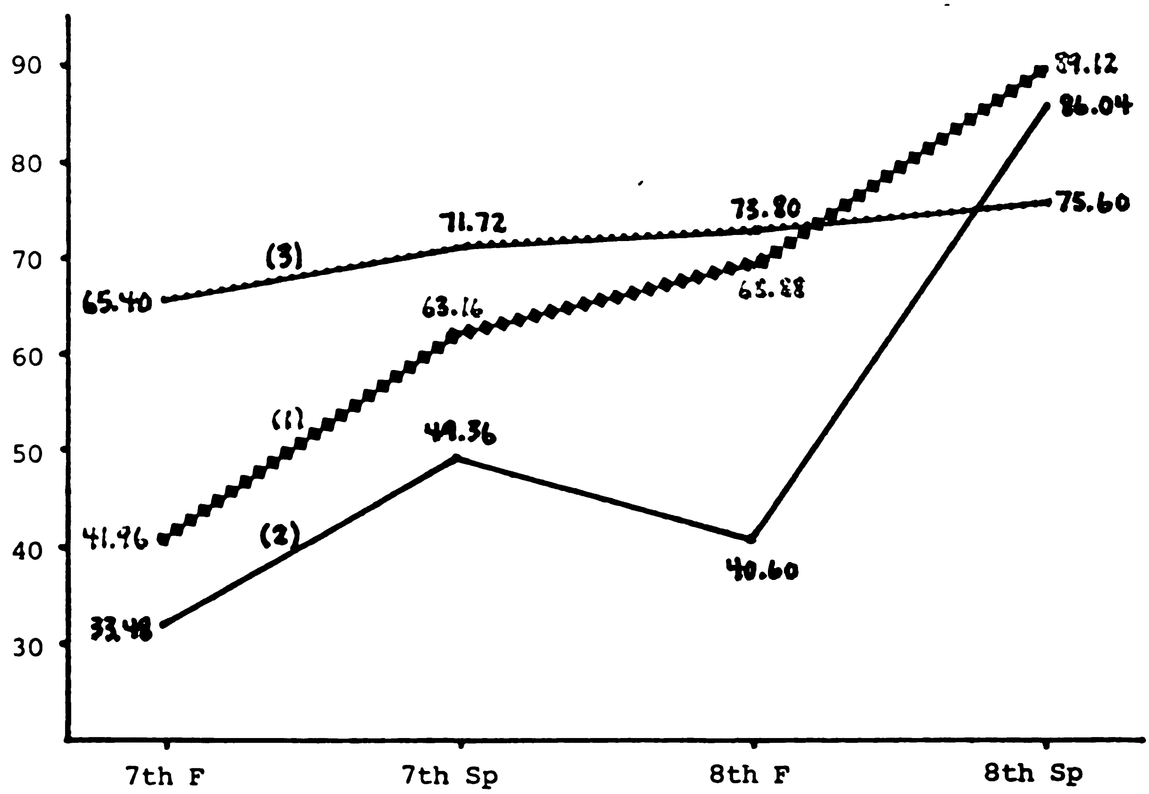
Variables	Univariate F	P less than
Chronol. Age	3.076	.0063
Height	.578	.7481
Weight	2.001	.0533
Sit-ups	5.222	.0001
Pull-ups	1.047	.3949
Stand Lg Jp	1.912	.0788
Vertical Jp	2.089	.0546
Shuttle run	2.159	.0471
Facial hair	1.492	.1818
Axillary hair	1.314	.2510
Pubic hair	.700	.6500
Penis dev.	1.189	.3119
Scrotum dev.	.726	.6287

The general F-ratio for the Mason GRADE by YEAR interaction was 2.3444. With a significance level less than .05, the analysis indicated that Mason's four levels of the GRADE factor and three levels of the year factor interacted significantly. However, because the four levels of the GRADE factor contained boys aged six months apart, the GRADE by YEAR interaction was probable. Therefore, the univariate F's and their significance levels became the interesting statistic.

Only the significance level for sit-ups was below .004. Figure L-3 graphically illustrates the interaction in sit-ups. The second year test period 8th F boys having a lower score than the 7th Sp boys and the third year test period 7th F boys and 8th Sp boys having mean scores that were extremely high and low, respectively, caused the significant differences in sit-ups.

Figure L-3

The Mason data's sit-ups mean values GRADE by YEAR interaction



Contrast #3 (SCHOOL by GRADE by YEAR)

The forty subjects for the East Lansing cells and the twenty-five subjects for the Mason cells fulfilled the proportional cells requirement. Tables L-30 through L-42 provide the East Lansing and Mason mean values for their proportional cell samples.

Table L-31

East Lansing and Mason data's chronological age mean values in their proportional cell samples

YEAR	7th F EL-M	7th Sp EL-M	8th F EL-M	8th Sp EL-M
1	150.01-150.15	155.61-156.36	162.96-164.68	167.19-173.02
2	149.41-150.66	155.74-157.50	163.08-161.22	168.43-167.12
3	149.71-152.86	155.56-.54.90	163.16-163.04	167.84-168.04

Table L-32

East Lansing and Mason data's height mean values in their proportional cell samples

YEAR	7th F EL-M	7th Sp EL-M	8th F EL-M	8th Sp EL-M
1	61-18-60.10	61.53-62.19	63.93-63.13	64.79-65.26
2	59.59-60.48	62.10-62.42	62.64-62.91	64.47-63.95
3	60.44-61.05	63.27-61.47	63.15-62.86	65.17-64.50

Table L-33

East Lansing and Mason data's weight mean values in their proportional cell samples

YEAR	7th F EL-M	7th Sp EL-M	8th F EL-M	8th Sp EL-M
1	97.12-95.32	109.85-110.12	118.82-123.84	120.43-123.56
2	95.65-100.84	106.76-119.72	110.14-113.34	127.41-114.22
3	101.45-108.54	111.20-105.12	111.39-112.48	123.40-124.76

Table L-34

East Lansing and Mason data's sit-ups mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	69.15-41.96	74.85-63.16	78.93-68.88	84.22-89.12
2	42.23-33.48	75.63-49.36	70.50-40.60	75.38-86.04
3	59.05-65.40	67.73-71.72	59.78-73.80	68.33-75.60

Table L-35

East Lansing and Mason data's pull-ups mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	3.18-2.88	2.20-3.84	3.50-3.00	3.88-6.36
2	2.30-3.60	2.93-2.36	3.30-3.00	2.43-5.44
3	2.88-2.28	2.98-3.32	3.63-2.44	4.05-4.52

Table L-36

East Lansing and Mason data's standing long jump mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	65.35-56.65	65.75-57.48	66.50-61.84	70.45-70.00
2	60.15-58.92	67.53-59.72	65.10-59.48	71.60-68.88
3	61.38-60.56	67.70-60.60	64.30-62.92	72.88-65.28

Table L-37

East Lansing and Mason data's vertical jump mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	14.73-11.40	13.93-12.00	15.58-11.44	16.58-14.60
2	13.13-11.32	14.85-11.60	14.13-12.40	16.05-13.96
3	14.05-12.12	13-73-12.68	14.60-13.60	16.13-13.72

Table L-38

East Lansing and Mason data's shuttle run mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	10.94-11.61	10.41-12.01	10.44-11.30	10.16-11.23
2	12.13-12.07	10.62-11.58	11.10-11.59	10.22-11.12
3	10.75-11.26	10.67-11.11	10.84-10.94	10.24-11.08

Table L-39

East Lansing and Mason data's facial hair mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	1.30-1.40	1.55-2.32	2.55-2.52	2.13-3.20
2	1.80-1.56	1.90-1.96	2.08-1.96	2.58-2.20
3	1.10-1.32	2.10-1.72	1.90-1.56	2.35-2.20

Table L-40

East Lansing and Mason data's axillary hair mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	1.08-1.40	1.63-2.24	2.40-2.84	2.43-2.92
2	1.25-1.60	1.68-2.16	1.88-2.00	2.65-2.12
3	1.33-1.56	1.85-1.88	2.43-1.84	2.68-2.40

Table L-41

East Lansing and Mason data's pubic hair mean values in their proportional cell samples

YEAR	7th F	7th Sp	8th F	8th Sp
	EL M	EL M	EL M	EL M
1	1.88-1.76	2.90-2.84	3.30-3.16	3.45-3.32
2	1.88-2.20	2.63-2.96	3.28-2.84	3.70-3.28
3	2.43-1.80	3.00-2.48	3.23-2.60	3.78-3.36

Table L-42

East Lansing and Mason data's penis development mean values in their proportional cell samples

YEAR	7th F		7th Sp		8th F		8th Sp	
	EL	M	EL	M	EL	M	EL	M
1	2.08-2.24		2.80-2.40		3.03-3.20		3.20-3.24	
2	2.15-2.20		2.63-2.68		3.03-2.44		3.63-3.24	
3	2.50-1.80		3.00-2.24		3.43-2.68		3.80-3.00	

Table L-43

East Lansing and Mason data's scrotum and testes development mean values in their proportional cell samples

YEAR	7th F		7th Sp		8th F		8th Sp	
	EL	M	EL	M	EL	M	EL	M
1	2.28-2.00		2.83-2.76		3.20-3.16		3.23-3.24	
2	2.10-2.28		2.75-2.52		3.05-2.64		3.70-3.04	
3	2.08-1.80		2.48-2.20		2.78-2.60		3.73-2.96	

Table L-44 provides the univariate F values and their associated significance levels for the East Lansing and Mason SCHOOL by GRADE by YEAR triple interaction.

Table L-44

The univariate F values and their associated significance levels for the East Lansing and Mason SCHOOL by GRADE by YEAR interaction

Variables	Univariate F	P less than
Chronol. Age	2.397	.0267
Height	1.276	.2657
Weight	1.746	.1077
Sit-ups	3.140	.0048
Pull-ups	1.654	.1295
Stand Lg Jp	2.914	.0082
Vertical Jp	2.377	.0279
Shuttle Run	1.864	.0845
Facial hair	2.571	.0180
Axillary hair	.897	.4963
Pubic hair	.813	.5598
Penis dev.	1.037	.3999
Scrotum dev.	1.127	.3447

The general F-ratio for the SCHOOL by GRADE by YEAR triple interaction was 2.2593. With a significance level less than .05, the analysis indicated the two SCHOOL levels across the four GRADE levels for the three YEAR levels interacted significantly. However, because the four levels of the GRADE factor contained boys aged six months apart, the SCHOOL by GRADE by YEAR interaction was probable. Therefore, the univariate F's and their significance levels became the interesting statistic.

Only the significance level for sit-ups was below .004. Figures L-4, L-5, and L-6 graphically illustrate the interaction in sit-ups.

Figure L-4

The East Lansing and Mason data's sit-ups mean values SCHOOL by GRADE interactions for the first year data collections

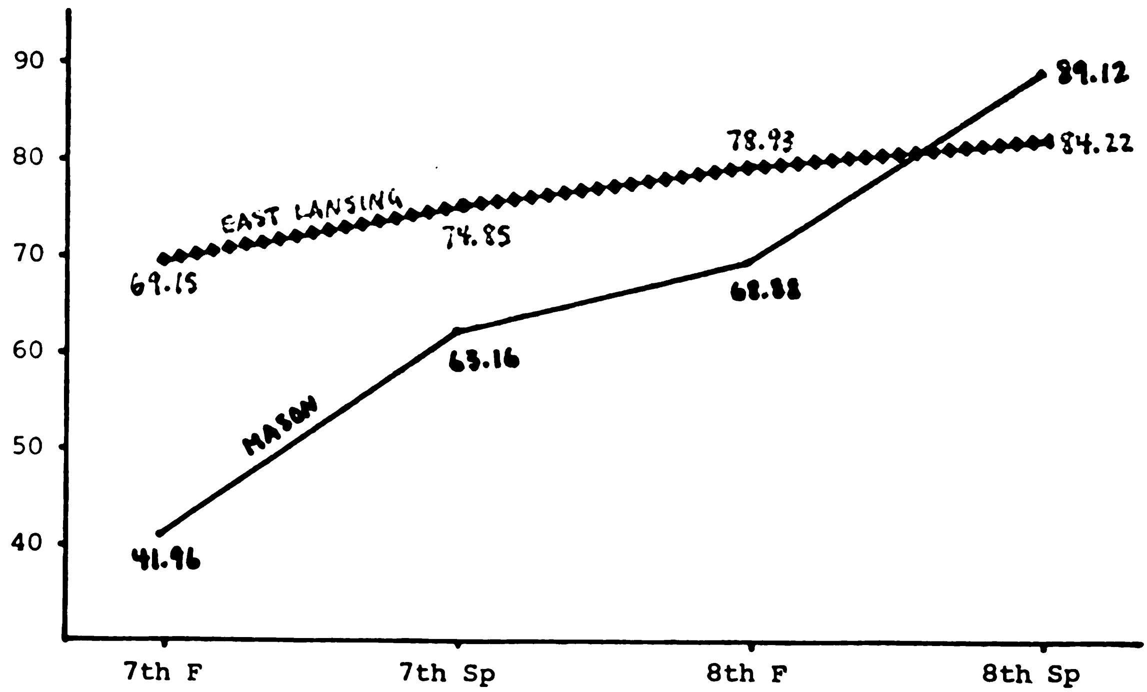


Figure L-5

The East Lansing and Mason data's sit-ups mean values
SCHOOL by GRADE interactions for the second year data collections

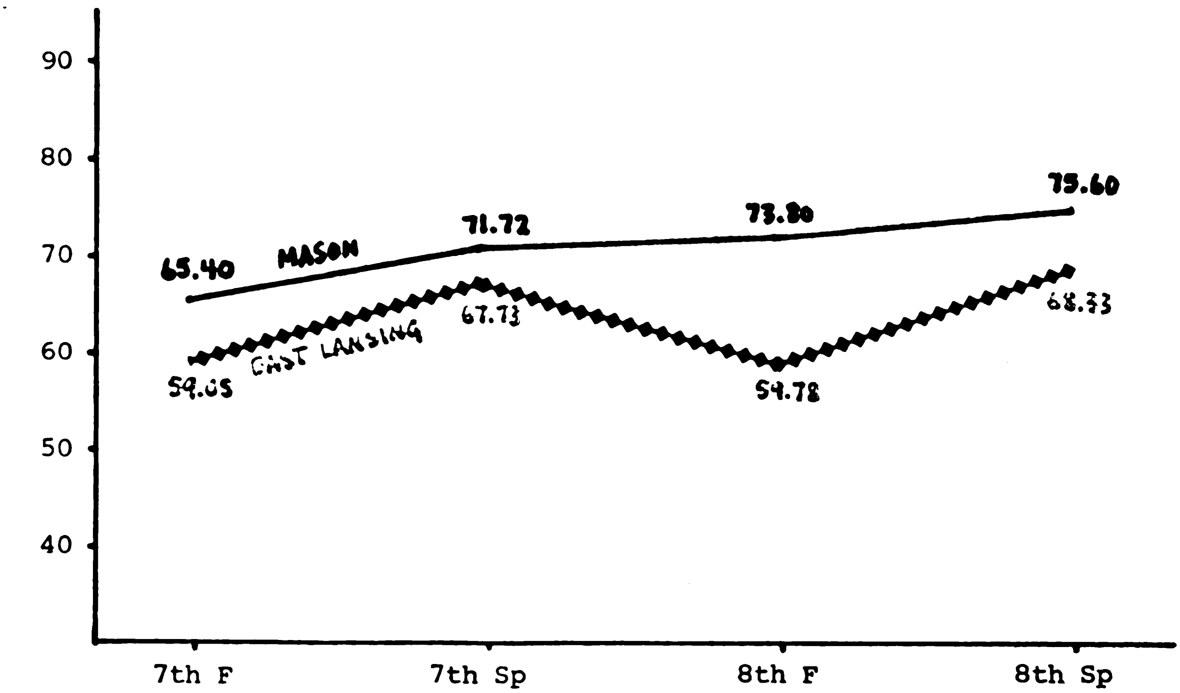
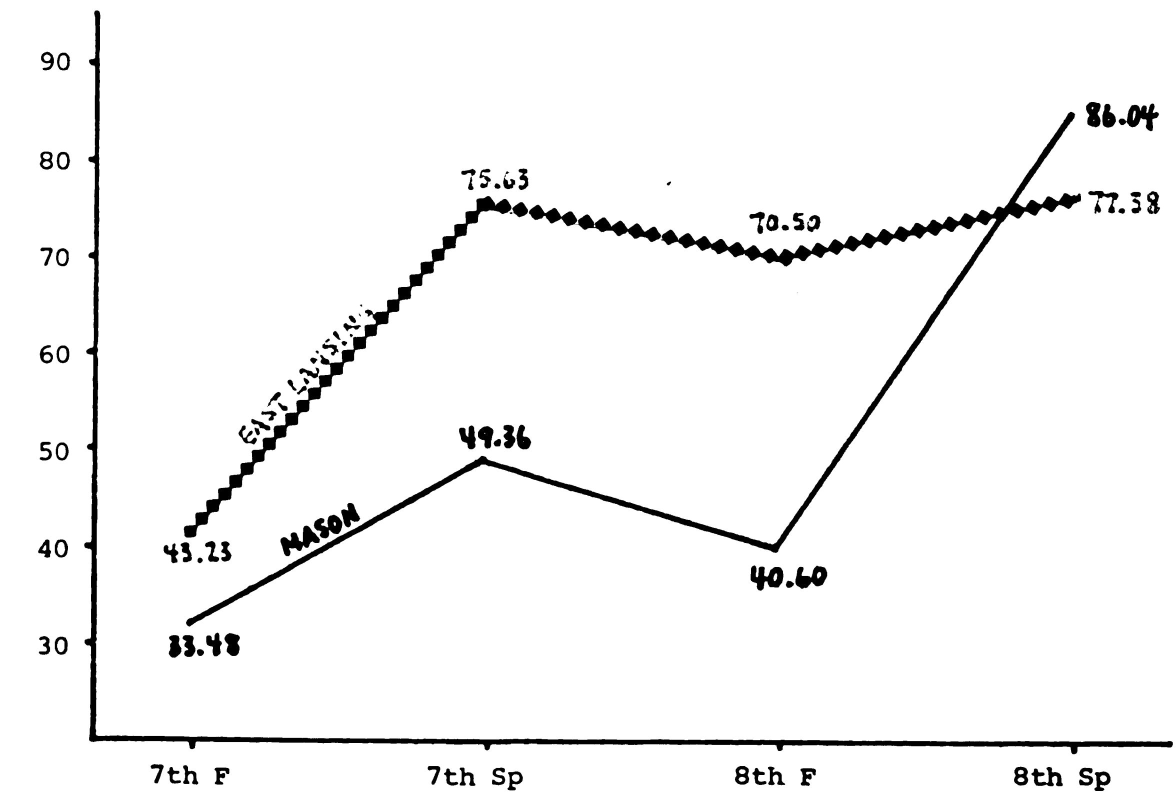


Figure L-6

The East Lansing and Mason data's sit-ups mean values
SCHOOL by GRADE interactions for the third year data collections



The Mason 8th Sp boys having a mean value higher than the East Lansing Sp boys may have contributed to the significant differences in sit-ups for the first year's data collections. The East Lansing and Mason 7th Sp boys having higher mean values than the 8th F boys contributed to the significant differences in sit-ups for the second year's collections. The East Lansing 7th Sp boys having an extremely high mean value and the East Lansing 8th F boys having an extremely low mean value contributed to the significant differences in sit-ups for the third year's data collection.

Another analysis that also examines the differences between the East Lansing and Mason subjects uses the least square estimates of the SCHOOL contrast. By doubling the standard error for each variable and adding and subtracting that value to and from its least square estimate, the 95 percent confidence interval for each variable is determined (Table L-45). If a confidence interval did contain zero, then the two school's

Table L-45

The least square estimates, standard errors, and confidence intervals of the thirteen variables

Variable	Least Square Estimates	Twiee its Standard Error	Confidence Interval	Contained Zero
Chronol. Age	-0.9033	2(0.4380)	-1.78 to -0.02	No
Height	0.1533	2(0.2561)	-0.37 to 0.67	Yes
Weight	-1.5204	2(1.7833)	-5.08 to 2.04	Yes
Sit-ups	5.7358	2(2.0489)	0.76 to 10.72	No
Pull-ups	-0.4763	2(0.2115)	-0.89 to -0.05	No
Stand Lg Jp	4.8629	2(0.6650)	3.52 to 6.20	No
Vertical Jp	2.2175	2(0.1936)	1.87 to 2.60	No
Shuttle run	-0.6915	2(0.0650)	-0.83 to -0.55	No
Facial hair	-0.0496	2(0.0740)	-0.18 to 0.10	Yes
Axillary hair	-0.1425	2(0.0925)	-0.32 to 0.04	Yes
Pubic hair	0.2354	2(0.0833)	0.08 to 0.40	No
Penis dev.	0.3242	2(0.0764)	0.16 to 0.48	No
Scrotum dev.	0.2479	2(0.0730)	0.11 to 0.39	No

data did not differ significantly. However, if a confidence interval did not contain zero, then the two school's data did differ significantly. If the data did differ between the two schools, then a judgement as to whether or not the difference was of practical significance must be made.

Because the contrast subtracted the Mason values from the East Lansing values, if a least square estimate was positive, then the East Lansing subjects had the higher value. The East Lansing subjects had higher values than the Mason subjects in height, sit-ups, standing long jump, vertical jump, pubic hair, penis development, and scrotum and testes development. Of those seven variables, the variables for which the 95 percent confidence intervals did not contain zero were sit-ups, standing long jump, vertical jump, pubic hair, penis development, and scrotum and testes development. Therefore, for sit-ups, standing long jump, vertical jump, pubic hair, penis development, and scrotum and testes development, the East Lansing subjects measured significantly higher than the Mason

The Mason subjects had higher values than the East Lansing subjects in chronological age, weight, pull-ups, shuttle run, facial hair, and axillary hair. Of those six variables, the variables for which the 95 percent confidence intervals did not contain zero were chronological age, pull-ups, and shuttle run. Therefore, for chronological age, pull-ups, and shuttle run, the Mason subjects measured significantly higher than the East Lansing subjects. The East Lansing and Mason subjects did not measure significantly different in height, weight, facial hair, and axillary hair.

Because the least square estimate analysis determined nine variables that were statistically different between the East Lansing and Mason contrast, the following suggestions about the practical significance of those differences are provided:

1) Because classification procedures for motor proficiency norms use yearly intervals, the .9 of a month advantage of the Mason subjects is not of practical significance.

2) Because a 5.74 sit-up difference does not exceed the difference between the mean values of the 7th F and 7th Sp boys or the 8th F and 8th Sp boys, the 5.74 sit-ups advantage of the East Lansing subjects is not of practical significance.

3) Because a .47 pull-ups difference exceeds the difference between the mean values of 7th F and 8th F boys, the .47 pull-ups advantage of the Mason subjects is of practical significance.

4) Because a 4.86 standing long jump difference exceeds the difference between the mean values of the 7th F and 7th Sp boys, the 4.86 standing long jump advantage of the East Lansing subjects is of practical significance.

5) Because a 2.22 vertical jump difference exceeds the difference between the mean values of the 7th F and 8th F boys, the 2.22 vertical jump advantage of the East Lansing subjects is of practical significance.

6) Because a .69 shuttle run difference exceeds the difference between the mean values of the 7th F and 8th F boys, the .69 advantage of the East Lansing subjects is of practical significance.

7) Because a .24 pubic hair difference does not exceed the difference between the mean values of the 7th F and 7th Sp boys or the 8th F and 8th Sp boys, the .24 pubic hair advantage of the Mason subjects if not of practical significance.

8) Because a .32 penis development difference does not exceed the difference between the mean values of the 7th F and 7th Sp boys or the 8th F and 8th Sp boys, the .32 penis development advantage of the Mason

PLEASE NOTE:

Page 257 is lacking in number only. No text is missing. Filmed as received.

UNIVERSITY MICROFILMS.

subjects is not of practical significance.

9) Because a .25 scrotum and testes difference does not exceed the difference between the mean values of the 7th F and 7th Sp boys or the 8th F and 8th Sp boys, the .25 scrotum and testes development advantage of the Mason subjects is not of practical significance.

To summarize, the purpose of the three contrast analyses and the least square estimates analysis was to examine the East Lansing and Mason data to determine whether or not by combining those data, the statistical analysis had a proper sample. The least square estimates analysis demonstrated that the motor proficiency tests of pull-ups, standing long jump, and shuttle run were of practical difference. However, the three-way interaction analysis demonstrated that only the sit-ups difference was statistically significant. Therefore, with some concern for the sit-ups test, the author concluded that combining the East Lansing and Mason data does provide a proper sample.

APPENDIX M

THE PROPORTION OF THIS SAMPLE'S 1578 BOYS WHO REACHED THE FIVE STAGES OF THE FIVE SEXUAL CHARACTERISTICS

Table M-1

The observed proportions for facial hair

AGES (mo.)	N	STAGES				
		1	2	3	4	5
132	15	100.0	0.0	0.0	0.0	0.0
138	45	71.1	26.7	2.2	0.0	0.0
144	85	74.1	20.0	5.9	0.0	0.0
150	234	59.4	30.8	7.7	1.7	0.4
156	283	46.6	36.0	12.4	3.9	1.1
162	284	37.7	29.9	21.5	9.5	1.4
168	248	28.2	27.4	25.8	14.1	4.4
174	159	15.1	21.4	27.7	28.9	6.9
180	122	9.8	10.7	23.0	40.2	16.4
186	79	2.5	6.3	19.0	46.8	25.3
192	14	0.0	0.0	7.1	57.1	35.7

Table M-2

The observed proportions for axillary hair

AGES (mo.)	N	STAGES				
		1	2	3	4	5
132	15	100.0	0.0	0.0	0.0	0.0
138	45	91.1	6.7	2.2	0.0	0.0
144	85	91.8	0.0	3.5	4.7	0.0
150	234	72.6	14.5	4.3	7.3	1.3
156	283	61.1	18.0	10.6	7.8	2.5
162	284	45.1	21.1	12.7	15.8	5.3
168	248	35.1	20.6	10.9	22.2	11.3
174	159	21.4	24.5	16.4	23.3	14.5
180	122	14.8	12.3	9.0	38.5	25.4
186	79	5.1	16.5	6.3	35.4	36.7
192	14	0.0	7.1	7.1	35.7	50.0

Table M-3

The observed proportions for pubic hair

AGES (mo.)	N	STAGES				
		1	2	3	4	5
132	15	80.0	13.3	6.7	0.0	0.0
138	45	60.0	20.0	17.8	2.2	0.0
144	85	58.8	24.7	7.1	9.4	0.0
150	234	32.9	30.3	17.5	17.9	1.3
156	283	20.1	30.7	26.9	17.7	4.6
162	284	5.6	24.3	28.2	32.7	9.2
168	248	9.3	14.5	22.6	35.5	18.1
174	159	4.4	7.5	30.8	35.2	22.0
180	122	2.5	6.6	17.2	37.7	36.1
186	79	2.5	1.3	8.9	36.7	50.6
192	14	7.1	0.0	0.0	35.7	57.1

Table M-4

The observed proportions for penis development

AGES(mo.)	N	STAGES 1	2	3	4	5
132	15	40.0	53.3	6.7	0.0	0.0
138	45	35.6	51.1	11.1	2.2	0.0
144	85	36.5	41.2	16.5	5.9	0.0
150	234	18.4	42.3	27.8	9.8	1.7
156	283	18.4	31.8	33.6	12.7	3.5
162	284	8.1	23.6	33.8	28.5	6.0
168	248	9.7	16.9	27.4	32.3	13.7
174	159	4.4	11.9	23.3	43.4	17.0
180	122	4.9	11.5	9.8	45.9	27.9
186	79	2.5	3.8	12.7	50.6	30.4
192	14	0.0	7.1	21.4	28.6	42.9

Table M-5

The observed proportions for scrotum and testes development

AGES(mo.)	N	STAGES 1	2	3	4	5
132	15	40.0	60.0	0.0	0.0	0.0
138	45	33.3	53.3	13.3	0.0	0.0
144	85	40.0	44.7	11.8	3.5	0.0
150	234	15.8	51.7	25.2	6.4	0.9
156	283	14.8	44.2	25.8	12.4	2.8
162	284	5.6	29.6	32.7	26.1	6.0
168	248	5.2	19.4	30.6	31.9	12.9
174	159	3.8	18.9	28.3	37.1	11.9
180	122	0.8	17.2	17.2	45.1	19.7
186	79	2.5	5.1	17.7	48.1	26.6
192	14	7.1	7.1	28.6	35.7	21.4

Figures M-1 through 5 graphically illustrate the proportion values contained in Tables M-1 through M-5.

Because the observed proportion figures for the five sexual characteristics were somewhat irregular, especially at the low and high ages, the author used the Quantal Response for Random Predictor Variables formula to calculate the predicted proportion values contained in Table M-6 through M-10 (Schmidt and McSweeney, 1977). Figures 6 through 10 graphically illustrate those proportion values.

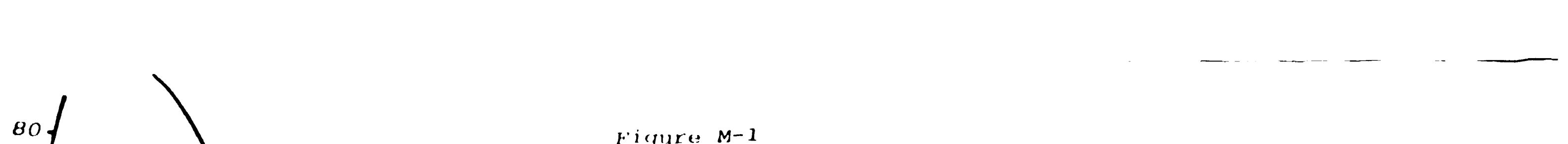
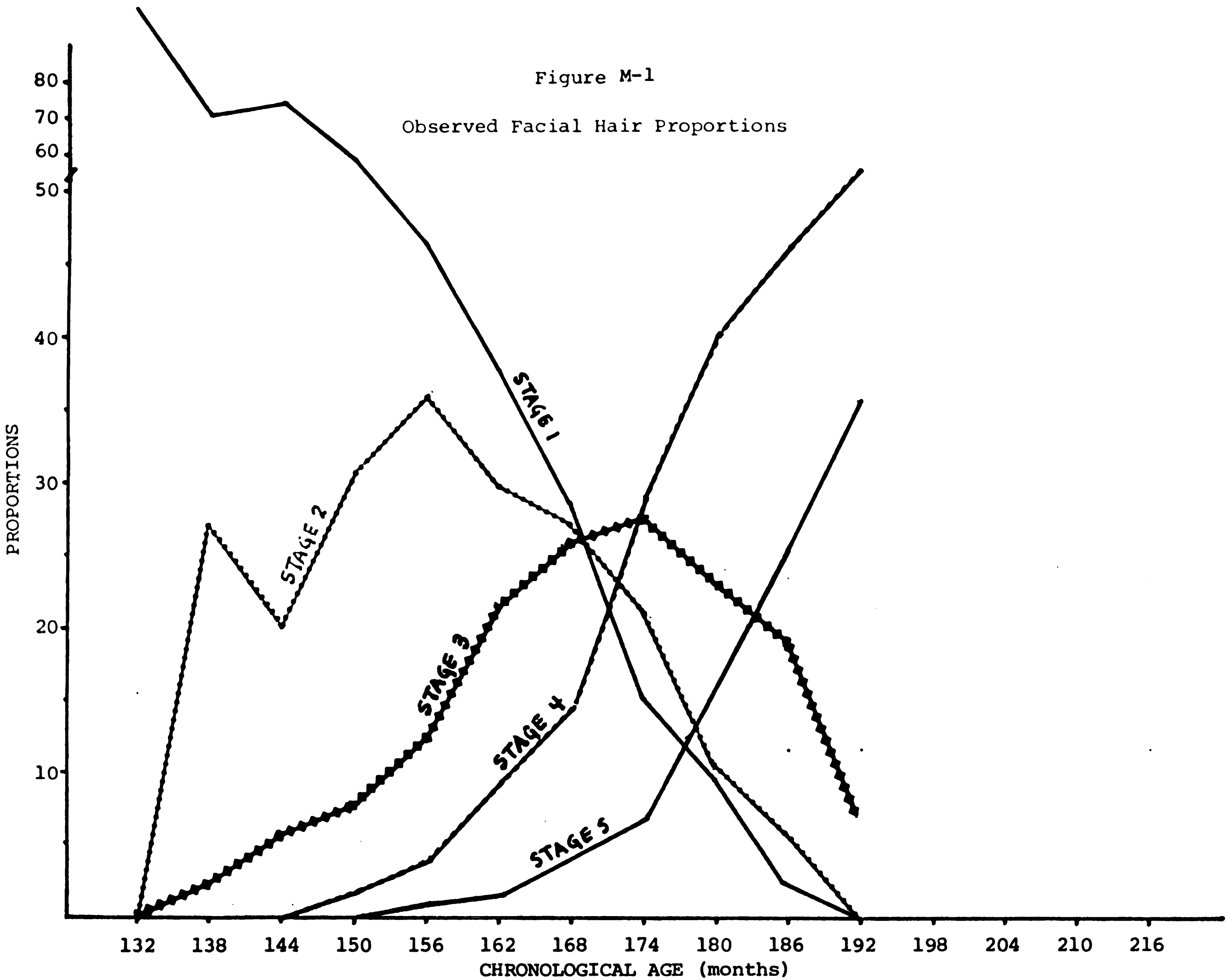


Figure M-1

Figure M-1

Observed Facial Hair Proportions




80

Figure M-2

Observed Axillary Hair Proportion

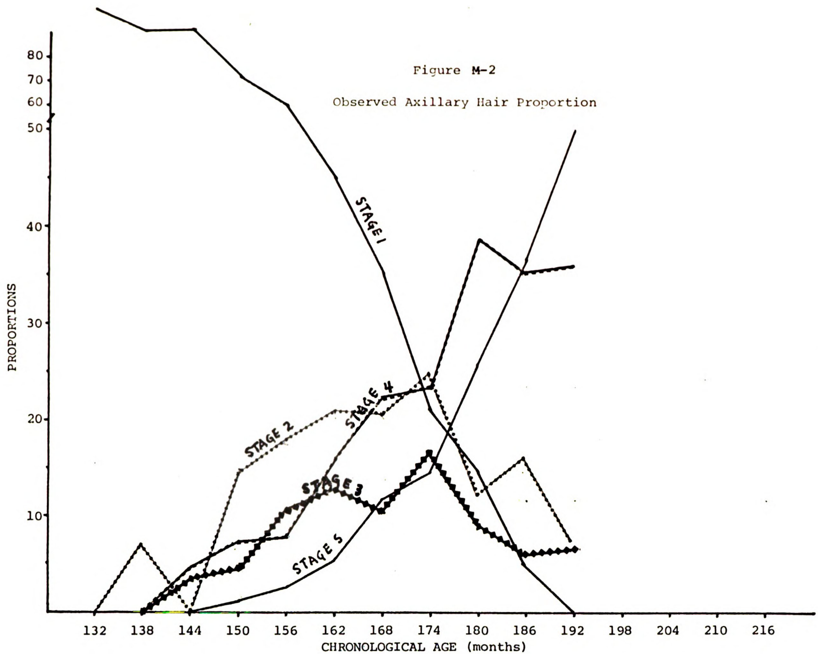


Figure M-3

80

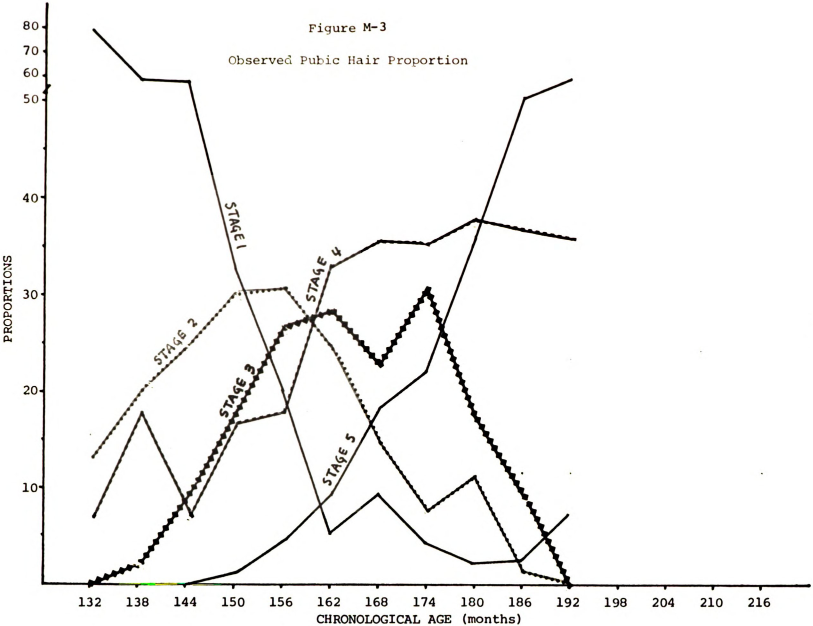


Figure M-3

Observed Pubic Hair Proportion

Figure M-4

Observed Penis Development Proportions

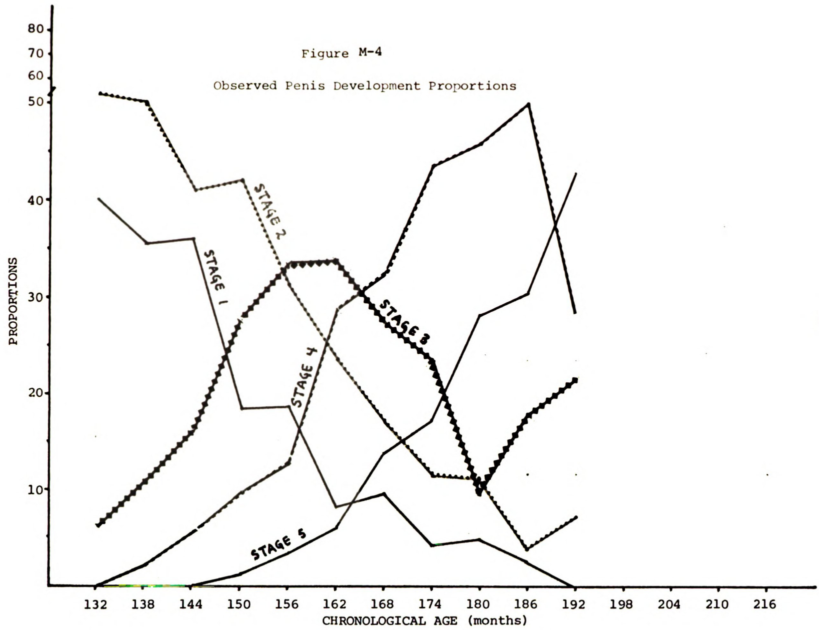


Figure M-5

Observed Scrotum and Testes Development Proportions

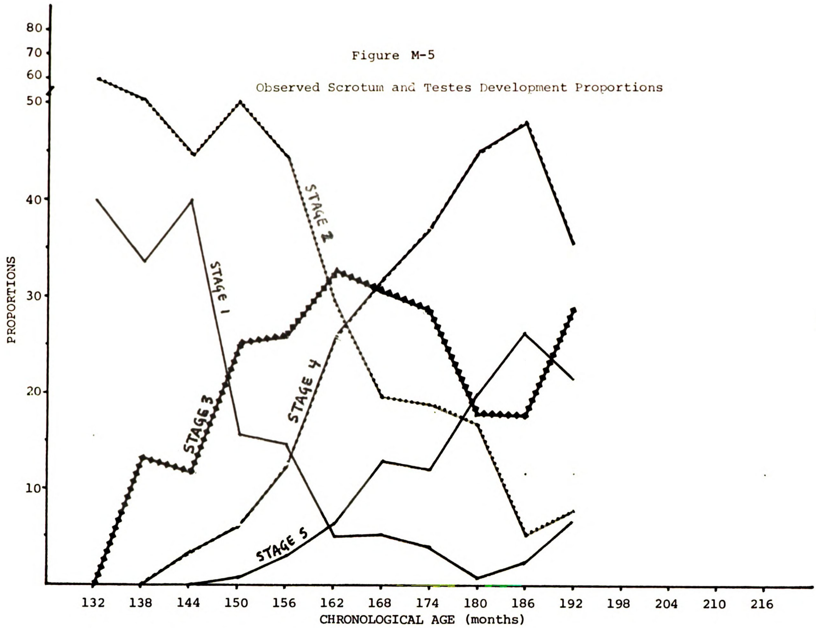


Table M-6

The predicted proportions for facial hair

AGES (mo.)	STAGES				
	1	2	3	4	5
132	80.1	18.3	1.3	0.1	0.0
138	75.0	22.2	2.7	0.2	0.0
144	68.6	26.2	4.8	0.5	0.1
150	60.4	29.9	8.2	1.3	0.2
156	50.4	32.3	13.2	3.4	0.7
162	38.5	31.9	19.6	8.1	1.8
168	25.7	27.6	25.4	16.8	4.6
174	14.4	19.9	27.4	28.9	9.4
180	6.6	11.9	24.4	41.2	15.9
186	2.6	6.0	18.6	49.8	23.0
192	0.9	2.7	12.3	54.0	29.7
198	0.3	1.2	8.0	54.7	35.8
204	0.1	0.5	4.8	53.1	41.4
210	0.0	0.2	2.9	50.2	46.7
216	0.0	0.0	1.7	46.4	51.8

Table M-7

The predicted proportions for axillary hair

AGES (mo.)	STAGES				
	1	2	3	4	5
132	93.4	4.3	1.5	0.7	0.1
138	89.5	6.4	2.5	1.4	0.1
144	83.7	9.3	4.0	2.7	0.4
150	74.3	13.0	6.0	5.2	1.4
156	62.8	16.9	8.5	9.3	2.3
162	48.3	20.2	11.1	15.3	5.1
168	33.2	21.5	12.9	22.4	10.0
174	20.1	20.2	13.2	29.1	17.3
180	10.9	16.9	12.0	33.5	26.7
186	4.3	12.8	9.9	35.6	37.1
192	2.4	8.9	7.6	33.4	47.7
198	1.0	5.9	5.4	30.2	57.6
204	0.3	3.7	3.7	26.0	66.3
210	0.1	2.2	2.4	21.6	73.6
216	0.0	1.3	1.6	17.4	79.6

Table M-8

The predicted proportions for pubic hair

AGES (mo.)	STAGES 1	2	3	4	5
132	66.6	24.2	6.8	2.3	0.1
138	57.1	28.1	10.0	4.5	0.2
144	45.5	30.6	14.9	9.1	0.8
150	31.4	32.5	20.2	14.2	1.7
156	21.9	27.5	24.7	21.8	4.0
162	12.9	22.3	26.7	29.7	8.4
168	6.1	15.6	24.6	35.9	17.8
174	3.1	9.9	22.5	38.7	25.8
180	1.5	5.7	17.5	37.5	37.7
186	0.5	3.0	9.5	32.9	54.0
192	0.2	1.4	8.0	26.7	63.7
198	0.1	0.6	4.9	20.3	74.1
204	0.0	0.3	2.8	14.7	82.2
210	0.0	0.1	1.6	10.3	88.0
216	0.0	0.0	0.9	7.0	92.1

Table M-9

The predicted proportions for penis development

AGES (mo.)	STAGES 1	2	3	4	5
132	35.9	47.4	15.6	0.8	0.1
138	31.8	45.9	19.6	2.5	0.3
144	27.3	44.9	23.8	3.9	0.7
150	22.4	39.4	27.8	8.7	1.7
156	17.4	33.6	30.5	15.0	3.6
162	12.3	26.2	30.8	23.8	6.9
168	7.9	18.5	28.5	33.3	11.9
174	4.5	11.6	23.1	42.8	18.0
180	2.3	6.3	17.2	48.9	25.4
186	1.1	3.4	11.0	52.1	32.4
192	0.5	1.7	7.3	51.4	39.1
198	0.2	0.8	4.5	49.3	45.2
204	0.1	0.4	2.8	45.7	50.9
210	0.0	0.2	1.6	42.1	56.2
216	0.0	0.1	0.9	37.9	61.2

Table M-10

The predicted proportions for scrotum and testes development

AGES (mo.)	STAGES				
	1	2	3	4	5
132	39.6	50.5	8.8	1.1	0.1
138	32.9	52.1	12.6	2.2	0.3
144	26.1	51.6	17.3	4.2	0.7
150	19.7	48.4	22.5	7.8	1.5
156	13.8	42.4	27.3	13.5	3.1
162	8.9	33.8	30.3	21.2	4.8
168	5.2	24.4	30.8	29.8	9.7
174	2.7	16.0	27.7	39.0	14.7
180	0.9	11.2	23.0	45.9	20.3
186	0.6	5.3	17.7	50.3	26.2
192	0.3	2.8	12.9	52.1	31.9
198	0.1	1.4	9.2	52.0	37.3
204	0.0	0.7	6.2	50.4	42.6
210	0.0	0.3	4.1	47.9	47.5
216	0.0	0.2	2.7	44.7	52.2

A comparison of the similarity of the observed proportion values with the predicted proportion models showed that the models fit the observed proportion model. Therefore, by examining Tables M-6 through M-10 and Figures M-6 through M-10, workers can determine the proportion of this 1,578 adolescent male sample that the raters rated as being in each of the five sexual characteristic's stage at 6 month intervals from 132 months to 216 months. For example, Table M-8 and Figure M-8 show that at age 162 months (13½ years), the raters rated; 12.9 percent of the 1,578 boys not yet started their pubic hair development, 22.3 percent as Stage 2 pubic hair, 26.7 percent as stage 3, 29.7 percent as stage 4, and 8.4 percent as having completed their pubic hair development.

Figure M-6

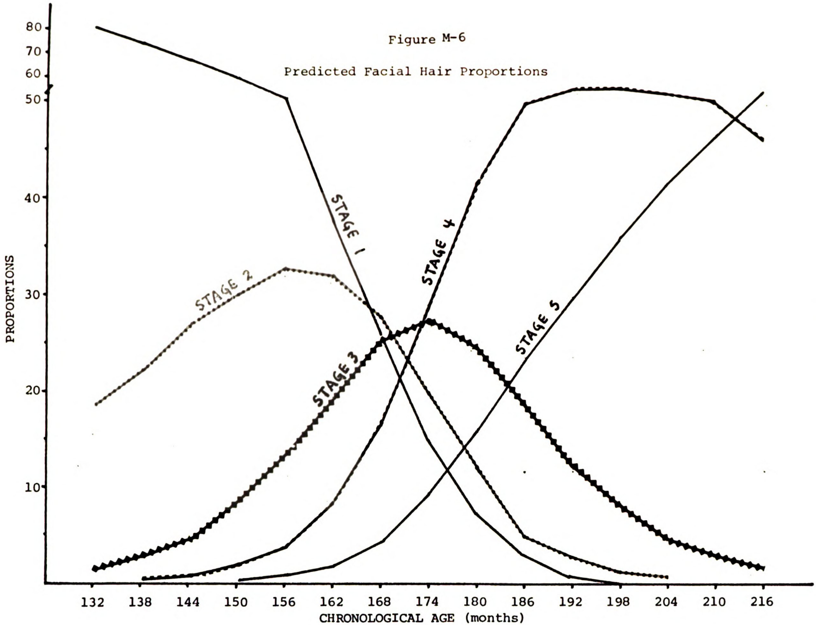


Figure M-6

Predicted Facial Hair Proportions

Figure M-7

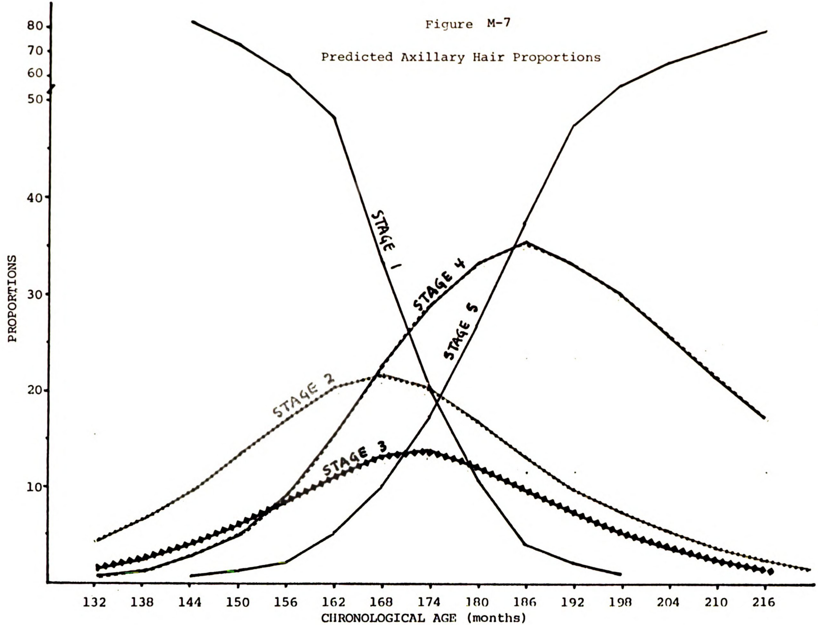


Figure M-7

Predicted Axillary Hair Proportions

Figure M-8

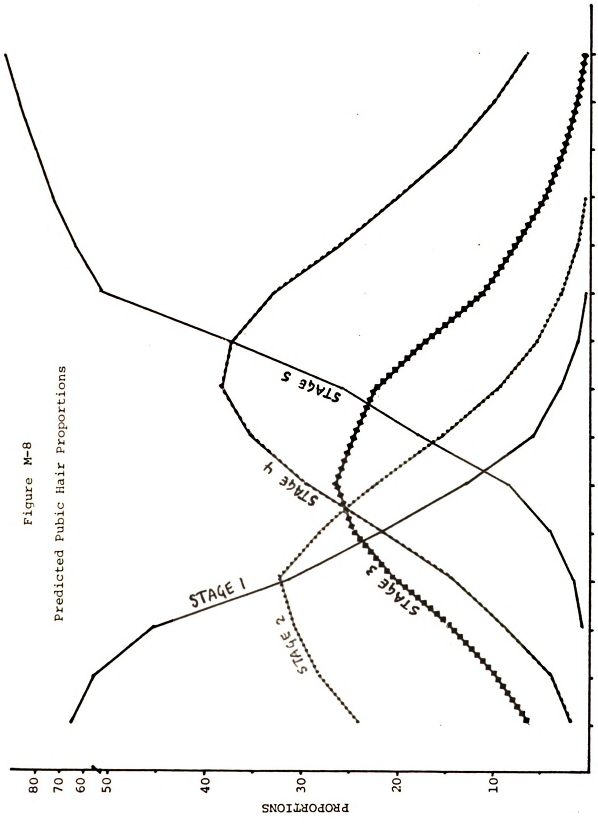


Figure M-8

Predicted Pubic Hair Proportions

Figure M-9

Predicted Penis Development Proportions

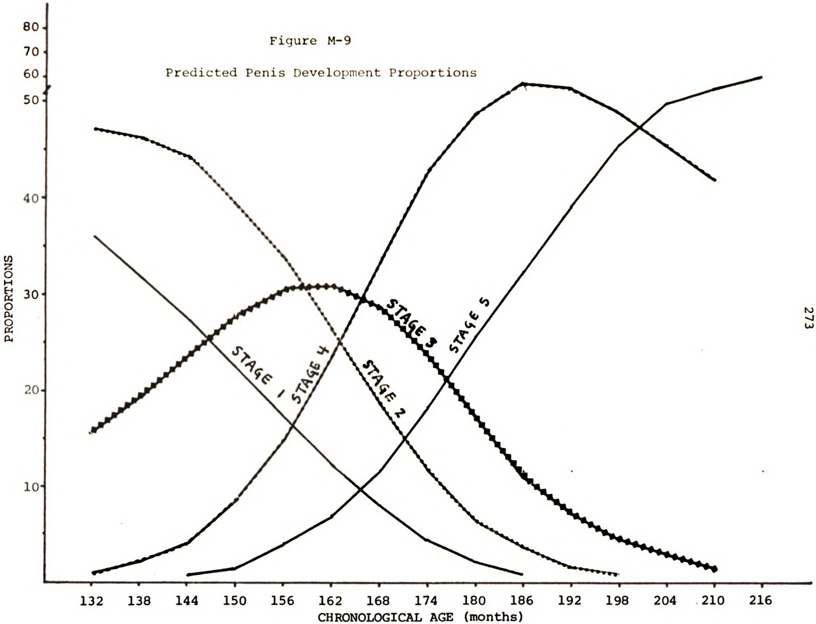
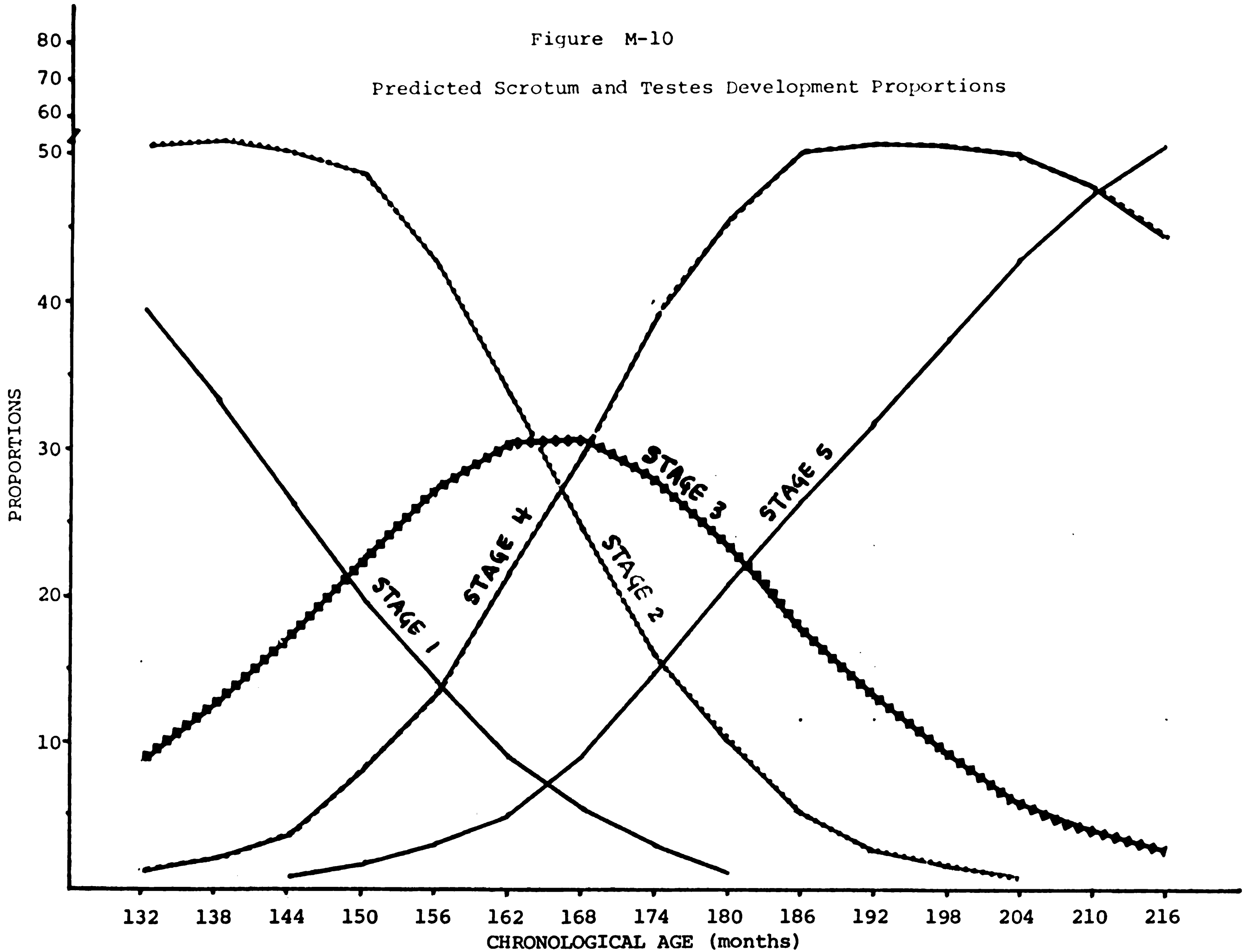


Figure M-10

Figure M-10

Predicted Scrotum and Testes Development Proportions



APPENDIX N

THE REGRESSION EQUATION FOR THE ESTIMATE OF SKELETAL AGE (ESA)

Before determining the regression equation for the estimate of skeletal age (ESA) that incorporates chronological age, height, weight, and SMIV, the correlation matrix for those four variables should be examined (Table N-1). The two anthropomorphic measures of height and

Table N-1

The sample chronological age, height, weight, and SMIV correlation matrix

	Chronological Age	Height	Weight	SMIV
Chronological Age	1.000			
Height	.698	1.000		
Weight	.388	.718	1.000	
SMIV	.560	.699	.537	1.000

weight correlated the highest (.718), while chronological age and weight correlated the lowest (.388). SMIV and height (.699) and chronological age and height (.698) correlated equally, but SMIV and chronological age correlated at only .560.

To obtain the regression equation for estimating skeletal age from chronological age, height, weight, and SMIV, the Multivariate Regression Analysis predicted the dependent variable (skeletal age) from those four independent variables. Table N-2 contains the raw regression coefficients and mean values that the Multivariate Regression Analysis determined for the ESA.

Table N-2

The raw regression coefficients and mean values that the Multivariate Regression analysis determined for the ESA regression equation

	Raw Regression Coefficients (β)	Mean Values (μ)
Chronological Age (CA)	.277	159.27
Height (H)	.212	63.01
Weight (W)	.220	108.30
SMIV	.594	160.57
ESA	---	160.53

The mean deviated (ESA) regression equation is:

$$ESA = \hat{\mu}(SA) + \hat{\beta}(CA)(X(CA)-\hat{\mu}(CA) + \hat{\beta}(H(X(H)-\hat{\mu}(H)) + \hat{\beta}(W)(X(W)) - \hat{\mu}(W)) + \hat{\beta}(SMIV)(X(SMIV) - \hat{\mu}(SMIV))$$

To convert this mean deviated regression equation to non-mean deviated, it is necessary to solve the following ESA $\hat{\beta}(0)$ equation;

$$ESA\ \hat{\beta}(0) = \hat{\mu}(SA) - \hat{\beta}(CA)\hat{\mu}(CA) - \hat{\beta}(H)\hat{\mu}(H) - \hat{\beta}(W)\hat{\mu}(W) - \hat{\beta}(SMIV)\hat{\mu}(SMIV)$$

After substituting the Table N-2 $\hat{\beta}$ and $\hat{\mu}$ values and solving this equation, ESA $\hat{\beta}(0)$ equalled -16.15.

To demonstrate the effects of different chronological age, height, weight and SMIV values on ESA, Table N-3 substitutes 144 and 168 months for chronological age, 60 and 72 inches for height, 100 and 150 pounds for weight, and 151.61 and 175.87 months for SMIV. The 151.61 and 175.87

Table N-3

Substituting selected chronological age, height, weight, and SMIV scores into ESA

$\hat{\beta}$(CA)(X(CA))	+	$\hat{\beta}$(H)(X(H))	+	$\hat{\beta}$(W)(X(W))	+	$\hat{\beta}$(SMIV)(X(SMIV))	+	$\hat{\beta}$(0)	=	ESA
.277 (144)	+	.212 (60)	+	.220 (100)	+	.594 (151.61)	+	(-16.15)	=	148.51
.277 (168)	+	.212 (60)	+	.220 (100)	+	.594 (151.61)	+	(-16.15)	=	155.16
.277 (144)	+	.212 (72)	+	.220 (100)	+	.594 (151.61)	+	(-16.15)	=	151.06
.277 (168)	+	.212 (72)	+	.220 (100)	+	.594 (151.61)	+	(-16.15)	=	157.71
.277 (144)	+	.212 (60)	+	.220 (150)	+	.594 (151.61)	+	(-16.15)	=	159.51
.277 (168)	+	.212 (60)	+	.220 (150)	+	.594 (151.61)	+	(-16.15)	=	166.16
.277 (144)	+	.212 (72)	+	.220 (150)	+	.594 (151.61)	+	(-16.15)	=	162.06
.277 (168)	+	.212 (72)	+	.220 (150)	+	.594 (151.61)	+	(-16.15)	=	168.71
.277 (144)	+	.212 (60)	+	.220 (100)	+	.594 (178.03)	+	(-16.15)	=	162.92
.277 (168)	+	.212 (60)	+	.220 (100)	+	.594 (175.87)	+	(-16.15)	=	169.57
.277 (144)	+	.212 (72)	+	.220 (100)	+	.594 (175.87)	+	(-16.15)	=	165.47
.277 (168)	+	.212 (72)	+	.220 (100)	+	.594 (175.87)	+	(-16.15)	=	172.12
.277 (144)	+	.212 (60)	+	.220 (150)	+	.594 (175.87)	+	(-16.15)	=	173.92
.277 (168)	+	.212 (60)	+	.220 (150)	+	.594 (175.87)	+	(-16.15)	=	180.57
.277 (144)	+	.212 (72)	+	.220 (150)	+	.594 (175.87)	+	(-16.15)	=	176.47
.277 (168)	+	.212 (72)	+	.220 (150)	+	.594 (175.87)	+	(-16.15)	=	183.12

months for SMIV represented the values when all five sexual characteristic ratings equalled 2 and 4, respectively. A 1 month increase in chronological age increased ESA .277 of a month, a 1 inch increase in height increased ESA .212 of a month, a 1 pound increase in weight increased ESA .220 of a month, and a 1 month increase in SMIV increased ESA .594 of a month.

As another way of examining the impact of those four independent variables on the dependent variable, provided that the range for chronological age is 60 months, the range for height is 20 inches, the range for weight is 100 pounds, and the range for SMIV is 50 months, by multiplying the beta weights for those variables times their ranges, the resulting value demonstrates the maximum impact potential for each variable. Table N-4 contains the beta weights, the ranges, and the maximum impact potentials for those four independent variables.

Table N-4

The beta weights, ranges, and maximum impact potentials for the four independent variables in the ESA regression equation

Variables	Beta Weights	x	Ranges	=	Max. Impact Potential
Chronological Age	.277	x	60	=	16.20 ESA months
Height	.212	x	20	=	4.24 ESA months
Weight	.220	x	100	=	22.00 ESA months
SMIV	.594	x	50	=	29.70 ESA months

SMIV and weight provide the highest impact potentials, with chronological age's impact about one-half of SMIV and height's impact very small. For chronological age to have the same impact as SMIV, the adolescent period would have to last almost 9 years. For height to have the same impact as SMIV, the boys would have to grow over 11½ feet during adolescence. And, for weight to have the same impact as SMIV, the boys would have to gain 135 pounds during adolescence.

To determine the precision that the analysis established the raw regression coefficients for ESA, you examine whether or not the confidence interval for each variable contains zero. A confidence interval equals the raw regression coefficient plus and minus twice its standard error. Table N-5 contains the raw regression coefficients and standard errors for the four variables in ESA.

Table N-5

The raw regression coefficients and standard errors for chronological age, height, weight, and SMIV

	Raw Regression Coefficients	Standard Errors
Chronological Age	.277	.142
Height	.212	.604
Weight	.220	.088
SMIV	.594	.142

Table N-6 contains the confidence intervals for those four variables.

Table N-6

The confidence intervals for chronological age, height, weight and SMIV

	Raw Regression Coefficient	+/-	twice its Standard Error	=	Confidence Interval
Chronological Age	.277	+ and -	2(.142)	=	(-.007 to .561)
Height	.212	+ and -	2(.604)	=	(-.996 to 1.42)
Weight	.220	+ and -	2(.088)	=	(.044 to .396)
SMIV	.594	+ and -	2(.142)	=	(.310 to .878)

The confidence intervals for chronological age (-.007 to .561) and height (-.999 and 1.420) both included zero. Therefore, for this sample the Multivariate Regression analysis did not establish the raw regression coefficients for chronological age and height very precisely.

Observed data contains measurement error. And, if one of the independent variables in a regression equation contains more measurement error than the other independent variables, the analysis' estimated raw regression coefficients can incorrectly attentuate the raw regression coefficient of the variable with the higher measurement error. Consequently, the interpretations of the impact of the individual independent variables can be misleading. Regression analyses on observed data cannot account for measurement error. In this analysis the chronological age, height, and weight independent variables should have far less measurement error in them than the SMIV independent variable. The solution is to use a Structural Multivariate Regression analysis to remodel the analysis. Structural Regression analyses take into account the measurement error contained in the SMIV variable (Lisrel, 1977).

The latent raw regression coefficients indicate whether or not the measurement error in SMIV distorted the ESA raw regression coefficients. While the Structural Multivariate Regression Analysis developed its own SMIV regression equation, it required that one of the sexual characteristic's beta weights equal 1.000. Therefore, the pubic hair characteristic beta weight was set equal to 1.000.

Table N-7 contains the parallel beta weights. While from the

Table N-7

The observed and structural ESA beta weights for chronological age, height, weight, and SMIV

	Chronological Age	Height	Weight	SMIV
Observed Beta Weights	.277	.212	.220	.594
Structural Beta Weights	.258	.120	.240	7.308

observed analysis to the structural analysis, chronological age's beta weight decreased slighly (.277 t0 .258) and weight's beta weight increased slightly (.220 to .240), height's beta weight decreased dramatically (.212 to .120). Because the SMIV scale for the observed analysis ranged from 1 to 5, those beta weights cannot be compared directly.

To compare those beta weights, a table similar to Table N-3 had to be developed for the structural values. From the discussion of the impact of the independent variables on the structural ESA, the relationship between the observed and structural SMIV becomes apparent.

To obtain the structural ESA regression equation, the Structural Multivariate Regression Analysis predicted the dependent variable (skeletal age) from the three independent variables of chronological age, height, and weight and the five sexual characteristics. However, whereas the analysis of the observed data provided beta weights for those eight variables individually, the structural analysis determined beta weights for chronological age, height, weight, and SMIV. Table N-8 contains the latent regression coefficients from the structural analysis and the mean values from the analysis on the observed data.

Table N-8

The latent regression coefficients from the structural analysis and the mean values from the analysis of the observed data

	Latent Regression Coefficients ($\hat{\beta}$)	Mean Values ($\hat{\mu}$)
Chronological Age	.258	159.27
Height	.120	63.01
Weight	.240	108.30
SMIV	7.308	cannot use
Skeletal Age	-----	160.53

Because the structural SMIV was in a different scale than the SMIV from the observed data, the structural ESA regression equation cannot use the observed $\hat{\mu}$(SMIV) value. Solving the following equation provided the $\hat{\mu}$(SMIV) value for the structural ESA regression equation.

$$\text{Structural } (\text{SMIV} = \{(\underline{\Lambda}' \Lambda)^{-1}\} \{(\Lambda'\hat{\mu})\}$$

Where $\underline{\Lambda}$ is the vector of the structural weight estimates for the five sexual characteristics and $\underline{\hat{\mu}}$ is the vector of the observed mean value estimates for the five sexual characteristics. Table N-9 contains the Λ vector and the $\hat{\mu}$ vector. After substituting the Table N-9 values and solving the equation, the structural $\hat{\mu}$(SMIV) equalled 3.03.

Table N-9

The Λ vector and the $\hat{\mu}$ vector

Λ vector	=	(0.689	0.719	1.000	1.053	0.769)
$\hat{\mu}$ vector	=	(2.04	1.86	2.95	3.02	2.98)

The mean deviated structural (S)ESA regression equation is;

$$(S)ESA = \hat{\mu}(SA) + \hat{\beta}(CA)(X(CA) - \hat{\mu}(CA) + \hat{\beta}(H)(X(H) - \hat{\mu}(H) + \hat{\beta}(W)(X(W) - \hat{\mu}(W)) + \hat{\beta}(SMIV)(X(SMIV) - \hat{\mu}(SMIV))$$

To convert this mean deviated equation to a non-mean deviated, it is necessary to solve the following equation;

$$(S)ESA\ \beta(0) = \hat{\mu}(SA) - \hat{\beta}(CA)\hat{\mu}(CA) - \hat{\beta}(H)\hat{\mu}(H) - \hat{\beta}(W)\hat{\mu}(W) - \hat{\beta}(SMIV)\hat{\mu}(SMIV)$$

After substituting the Table N-8 values and the calculated structural (SMIV) value and solving the equation, the ESA β(0) equalled 63.764.

To demonstrate the effects of increments of chronological age, height, weight and SMIV values on the (S)ESA values, Table N-10 substitutes 144 and 168 pounds for weight, and 2.00 and 4.00 for SMIV.

Table N-10

Substituting selected chronological age, height, weight, and SMIV scores into the (S)ESA regression equation

$\hat{\beta}$(CA)(X(CA))+$\hat{\beta}$(H)(X(H))+$\hat{\beta}$(W)(X(W))+$\hat{\beta}$(SMIV)(X(SMIV))+	β(0)	=	ESA
.258 (144) + .120 (60) + .240 (100) + 7.308 (2.00) +	63.764	=	146.73
.258 (168) + .120 (60) + .240 (100) + 7.308 (2.00) +	63.764	=	152.92
.258 (144) + .120 (72) + .240 (100) + 7.308 (2.00) +	63.764	=	148.17
.258 (168) + .120 (72) + .240 (100) + 7.308 (2.00) +	63.764	=	154.36
.258 (144) + .120 (60) + .240 (150) + 7.308 (2.00) +	63.764	=	158.73
.258 (168) + .120 (60) + .240 (150) + 7.308 (2.00) +	63.764	=	164.92
.258 (144) + .120 (72) + .240 (150) + 7.308 (2.00) +	63.764	=	160.17
.258 (168) + .120 (72) + .240 (150) + 7.308 (2.00) +	63.764	=	166.36
.258 (144) + .120 (60) + .240 (100) + 7.308 (4.00) +	63.764	=	161.35
.258 (168) + .120 (60) + ,240 (100) + 7.308 (4.00) +	63.764	=	167.54
.258 (144) + .120 (72) + .240 (100) + 7.308 (4.00) +	63.764	=	162.79
.258 (168) + .120 (72) + .240 (100) + 7.308 (4.00) +	63.764	=	168.98
.258 (144) + .120 (60) + .240 (150) + 7.308 (4.00) +	63.764	=	173.35
.258 (168) + .120 (60) + .240 (150) + 7.308 (4.00) +	63.764	=	179.54
.258 (144) + .120 (72) + .240 (150) + 7.308 (4.00) +	63.764	=	174.79
.258 (168) + .120 (72) + .240 (150) + 7.308 (4.00) +	63.764	=	180.98

As another way of examining the impact of those four independent variables on the dependent variable, provided that the range for chronological age is 60 months, the range for height is 20 inches, the range for weight is 100 pounds, and the range for SMIV is 4 levels, by multiplying the beta weights for those variables times their ranges, the resulting value demonstrates the maximum impact potential for each variable. Table N-11 contains the beta weights, ranges, and maximum impact potential for the four independent variables.

Table N-11

The beta weights, ranges, and maximum impact potentials for the four independent variables in the (S)ESA regression equation.

Variables	(S)ESA Beta Weights	x	Ranges	=	Max. Impact Poten.
Chronological Age	.258	x	60	=	15.48 (S)ESA months
Height	.120	x	20	=	2.40 (S)ESA months
Weight	.240	x	100	=	24.00 (S)ESA months
SMIV	7.308	x	4	=	29.33 (S)ESA months

SMIV and weight provide the highest impact potentials, with chronological age's impact about one-half of SMIV and the impact of height is minimal. For chronological age to have the same impact as SMIV, the adolescent period would have to last almost 9½ years. For height to have the same impact as SMIV, the boys would have to grow over 20 feet during adolescence. And, for weight to have the same impact as SMIV, the boys would have to gain almost 122 pounds during adolescence.

Table N-12 contains the ESA maximum impact potentials from the observed (O) and structural ESA maximum impact potentials. With the possible exception of height, the Structural Multivariate Regression analysis redistributed the beta weights very slightly.

Table N-12

The (O)ESA and (S)ESA maximum impact potentials

Variables	(O)ESA Maximum Impact Potentials	(S)ESA Maximum Impact Potentials
Chronological Age	16.20 (O)ESA months	15.48 (S)ESA months
Height	4.24 (O)ESA months	2.40 (S)ESA months
Weight	22.00 (O)ESA months	24.00 (S)ESA months
SMIV	29.70 (O)ESA months	29.23 (S)ESA months

Because the structural analysis eliminated the SMIV measurement error, the structural sample correlation matrix could be different from the observed sample correlation matrix. The structural analysis' PHI matrix was its sample variance matrix. Table N-13 contains the structural analysis' PHI matrix.

Table N-13

The structural analysis' PHI matrix

	Chronological Age	Height	Weight	SMIV
Chronological Age	174.713			
Height	40.267	19.056		
Weight	111.556	68.243	474.051	
SMIV	8.453	3.372	12.161	1.160

Dividing the off-diagonal variances by the product of the square roots of the diagonal variances provided the structural sample correlation matrix. Table N-14 contains the structural sample correlation matrix.

Table N-14

The structural sample correlation matrix

	Chronological Age	Height	Weight	SMIV
Chronological Age	1.000			
Height	.698	1.000		
Weight	.388	.718	1.000	
SMIV	.594 (.560)	.717 (699) -	.519 (.537)	1.000

In parenthesis beside the SMIV correlations are the comparable observed SMIV correlations. While from the observed analysis to the structural analysis the SMIV-chronological age correlation increased (.560 to .594) and the SMIV-height correlation increased (.699 to .717), the SMIV-weight correlation decreased (.537 to .519).

The structural analysis provided another statistical procedure for estimating the reliability of the five sexual characteristic ratings. Because the structural analysis used the average of the five rater's two assessments for each sexual characteristic as that characteristic's observed value, the following reliabilities represent the reliability of those averaged ratings.

The following variance (σ^2) equation provides the values that substitute into the structural reliability equations.

$$\sigma^2(X(i)) = \lambda^2(i)\sigma^2(\xi) + \sigma^2(e(i))$$

where the λ'x are the structural analysis' estimates of the sexual characteristic beta weights, the $\sigma^2(\xi)$ is the structural analysis' estimate of the SMIV variance, and the $\sigma^2(e)$'s are the structural analysis' estimates of the sexual characteristics measurement error and appear as the diagonal elements of the theta delta matrix. Table N-15 contains the $\sigma^2(X(i))$ equations for the five sexual characteristics and their solutions.

Table N-15

The $\sigma^2(X(i))$ equations for the five sexual characteristics and their solutions

	$\lambda^2(i)$		$\sigma^2(\xi)$		$\sigma^2(X(i))$		$\sigma^2(X(i))$
Facial Hair	.689	X	1.160	+	.390	=	.940
Axillary Hair	.719	X	1.160	+	.538	=	1.138
Pubic Hair	1.000	X	1.160	+	.159	=	1.319
Penis Dev.	1.000	X	1.160	+	.058	=	1.344
Scrotum Dev.	.769	X	1.160	+	.036	=	.722

The following reliability (ρ) equation provides the structural reliability values for the five sexual characteristics;

$$\rho(x(i)) = \lambda^2(i)\ \sigma^2(\xi) \div \sigma^2(x(i))$$

Table N-16 contains the $\rho(x(i))$ equations for the five sexual character-

Table N-16

The $\rho(x(i))$ equations for the five sexual characteristics and their solutions

	$\lambda^2(i)$	X	$\sigma^2(\xi)$	÷	$\sigma^2(x(i))$	=	$\rho(x(i))$
Facial Hair	.689	X	1.160	÷	.940	=	.586
Axillary Hair	.719	X	1.160	÷	1.138	=	.527
Pubic Hair	1.000	X	1.160	÷	1.319	=	.879
Penis Dev.	1.053	X	1.160	÷	1.344	=	.957
Scrotum Dev.	.769	X	1.160	÷	.722	=	.950

istics and their solutions. The averaged penis development ratings provided the highest reliabilities (.957), and the averaged axillary hair ratings provided the lowest reliability (.527). The sexual characteristics that provided the second, third, and fourth highest reliabilities were scrotum and testes development (.950), pubic hair (.879), and facial hair (.586).

APPENDIX O

AN ADDITIONAL CLASSIFICATION PROCEDURE THAT THE RESULTS SUGGESTED

The tremendous difference in the distribution of the subjects by ESA when compared to CA suggested another way of classifying the boys. By crossing CA with ESA, I determined the proportion of the boys in this sample who displayed accelerated, standard, and delayed biological maturation (Table O-1). Those subjects whose CA was higher than ESA, I

Table O-1

The proportion of the boys in this sample in each cell when CA is crossed with ESA

	11.0	12.0	13.0	14.0	15.0	16.0
11.0	17(43.6)	45(20.7)	10(1.8)			DELAYED
12.0	20(51.3)	104(47.9)	140(25.1)	31(6.5)	4(1.7)	
13.0	2(5.1)	50(23.0)	215(38.5)	109(22.9)	14(5.9)	
14.0		15(6.9)	150(26.9)	173(36.3)	47(19.9)	4(8.5)
15.0		3(1.4)	38(6.8)	136(28.6)	110(46.6)	16(34.0)
16.0			4(0.2)	24(5.0)	53(72.5)	21(24.7)
17.0	ACCELERATED		1(0.2)	3(0.6)	7(3.0)	5(10.6)

labeled delayed-maturers. Those subjects whose CA equalled ESA, I labeled standard-maturers. And, those subjects whose CA was lower than ESA, I labeled accelerated-maturers.

In Table O-2, I provided the percentage of accelerated, standard, and delayed-maturers.

Table O-2

The percentage of the accelerated, standard, and delayed-maturers across the six CA intervals

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	56.4	31.4	34.6	34.3	25.9	12.8
Standard	43.6	47.9	38.5	36.3	46.6	44.7
Delayed	----	20.7	26.9	29.4	27.5	42.5

Because the 12.0 through the 15.0 age range contains the greatest variability in adolescent male biological maturation, I will consider those ages of prime importance in the following analyses. The percent of standard-maturers decreased from 48% to 36% from 12.0 to 14.0 but returned to 42% at 15.0. The percent of the accelerated-maturers increased from 31% at 12.0 to a high of 35% at 13.0 and decreased to 26% at 15.0. The percent of delayed-maturers increased from 21% at 12.0 to 29% at 14.0 and settled back to 28% at 15.0.

To verify this new classification procedure, I needed to determine the variances for the motor proficiency tests (Table O-3).

Table O-3

The variances for the motor proficiency tests when classifying by CA, ESA, and the accelerated, standard, and delayed classifications

	Sit-ups	Pull-ups	Standing Long Jump	Vertical Jump	Shuttle Run
CA	763.55	9.97	94.97	8.82	0.86
ESA	800.28	10.33	93.82	8.66	0.87
Acclerated	739.72	9.84	103.94	10.30	0.91
Standard	757.98	9.78	89.25	7.32	0.78
Delayed	802.07	9.27	71.49	6.78	0.81

For the sit-ups test, the accelerated and standard categories both displayed slightly lower variances than either CA or ESA. The accelerated-maturers had more homogeneity than either the standard or delayed-maturers. The reason for this finding might be that more of the accelerated-maturers attained the maximum number of sit-ups. If that hypothesis is correct, then the accelerated-maturers will be more homogeneous than the standard or delayed-maturers during the early adolescent years.

For the pull-ups test, the accelerated, standard, and delayed categories all had lower variances than either CA or ESA. The delayed-maturers had more homogeneity than the other categories. The reason for that finding might be that the delayed-maturers cannot do very many pull-ups, therefore, they score similarly. Conversely, some accelerated-maturers can do many pull-ups, hence, they display less homogeneity. If that hypothesis is correct, then the variances for the delayed-maturers should increase across CA.

For the standing long jump test, the standard and delayed categories both displayed lower variances than either CA or ESA. The delayed category had a 24% decrease in variance from the ESA variance. The delayed-maturers again had a lower variance than either of the other two categories. The reason for this finding might be that a few highly-skilled and motivated accelerated-maturers score far better than their peers. If this hypothesis is correct, then the variances for the accelerated-maturers will increase across CA.

The results of the vertical jump test parallel the standing long jump results.

For the shuttle-run test, the standard and delayed categories both displayed slightly lower variances than either CA or ESA. The standard-maturers had slightly more homogeneity than the delayed-maturers and 14% more homogeneity than the accelerated-maturers. This finding is a puzzle. Apparently, biological maturation does not influence the shuttle run tests in the same manner as it effects the other tests. Because the shuttle run test measures agility rather than strength or power, the hypothesis might be that biological maturation does not effect agility.

In Tables O-4 through O-8, I have presented the variances for the five motor proficiency tests across the six CA intervals. The asterisk indicates that the variance was lower than either the CA or ESA variances.

Table O-4

The sit-ups variances across the six CA intervals for the accelerated, standard, and delayed-maturers

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	7.14*	6.00*	7.25*	13.55	15.47	30.20
Standard	9.10*	5.19*	8.80*	8.86*	15.89	4.38*
Delayed	----	6.59*	7.57*	9.49*	13.55	6.48*

Table O-5

The pull-ups variances across the six CA intervals for the accelerated, standard, and delayed-maturers

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	7.14*	6.00*	7.25*	13.55	15.47	30.20
Standard	9.10*	5.19*	8.80*	8.86*	15.89	4.38*
Delayed	----	6.59*	7.57*	9.49*	13.55	6.48*

Table O-6

The standing long jump variances across the six CA intervals for the accelerated, standard, and delayed maturers

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	95.64	76.97*	89.12*	125.44	157.41	288.50
Standard	20.03*	65.66*	72.01*	105.54	115.91	173.93
Delayed	-----	61.22*	67.42*	70.19*	88.05*	92.83*

Table O-7

The vertical jump variances across the six CA intervals for the accelerated, standard, and delayed-maturers

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	7.92*	8.19*	8.80*	12.55	12.25	17.30
Standard	4.01*	4.88*	6.85*	11.09	8.52*	9.53
Delayed	----	5.89*	5.29*	6.83*	11.70	7.47

Table O-8

The shuttle run variances across the six CA intervals for the accelerated, standard, and delayed-maturers

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	1.32	1.11	0.90	0.82*	1.01	0.90
Standard	0.50*	0.73*	0.83*	0.90	0.63*	0.55*
Delayed	----	0.83*	0.93	0.77*	0.72*	0.69*

For the sit-ups test, I hypothesized that the accelerated-maturers would have lower variances during the early years than the standard and delayed-maturers. Not only does that hypothesis stand up, but at 14.0, 15.0, and 16.0, the variances for all of the categories decreased dramatically. The variance for the 15.0 standard-maturers reduced 41% from the CA variance.

For the pull-ups test, I hypothesized that the delayed-maturers had lower variances because they could not do many pull-ups. Therefore, as the CA intervals increased, the variances for the delayed-maturers should increase. From 12.0 to 15.0, the variance for the delayed-maturers increased from 6.59 to 13.55. The pull-ups results provides no pattern. At 12.0 and 14.0, the standard-maturers have the lowest variances. At 13.0, the accelerated-maturers have the lowest variance. And, at 15.0, the delayed-maturers have the lowest variance. However, we should note that with the one exception of the accelerated-maturers at 14.0, all of the categories for the 12.0, 13.0, and 14.0 CA levels have lower variances than either CA or ESA. And, the variance for the 12.0 standard-maturers is 38% less than the CA variance.

For the standing long jump test, I hypothesized that the highly-skilled and motivated accelerated-maturers created the higher variances for the accelerated-maturers category. Therefore, as the CA intervals increased, the variances for the accelerated-maturers should increase. From 12.0 to 15.0, the variance for the accelerated-maturers increased from 76.97 to 157.41. There are two very clear patterns in the standing long jump results. (1) As CA increases, the variances for all three categories increase. (2) At every CA level, the accelerated-maturers show the least homogeneity and the delayed-maturers show the most homogeneity. Also, we should note that with the two exceptions of the accelerated and standard-maturers at 14.0, all of the categories for the 12.0, 13.0, and 14.0 CA levels have lower variances than either CA or ESA.

For the vertical jump test, I hypothesized the same situation as for the standing long jump test. And indeed, from 12.0 to 15.0, the variances for the accelerated-maturers increased from 8.19 to 12.25 with a high of 12.55 at 14.0. Unlike the standing long jump finding, however, every category did not increase as CA increased. The accelerated and standard-maturers increased from 12.0 to 14.0, but they both decreased from 14.0 to 15.0. Also, unlike the standing long jump findings, the relationship between the three categories did not remain constant across the CA intervals. At 12.0 and 15.0 the standard-maturers had the most homogeneity. At 13.0 and 14.0, the delayed-maturers had the most homogeneity. Also, we should note that with the three exceptions of the accelerated-maturers at 13.0 and 14.0 and the standard-maturers at 14.0, all of the categories for the 12.0, 13.0, and 14.0 CA levels have lower variances than either CA or ESA.

The lower variances for the 12.0, 13.0, and 14.0 CA levels for pull-ups, standing long jump, and vertical jump indicate that this method classifies adolescent males during the ages of the greatest variability in their biological maturation better than CA or ESA, and, some of the reductions in variance are considerable. For example, the 38% reduction for standard-maturers at 12.0 for pull-ups, the 35% reduction for delayed-maturers at 12.0 for standing long jump, and the 44% reduction for standard-maturers at 12.0 for vertical jump.

For the shuttle run test, the results were puzzling. The variances of the accelerated-maturers decreased from 12.0 to 14.0 and increased from 14.0 to 15.0. The variances of the standard-maturers increased from 12.0 to 14.0 and decreased from 14.0 to 15.0. The variances of the delayed-maturers increased from 12.0 to 13.0 and decreased from 13.0 to 15.0. The standard-maturers had the most homogeneity at 12.0, 13.0, and 15.0 with the delayed-maturers the most homogeneous at 14.0. Three of the four variances from 12.0 to 15.0 were lower than either CA or ESA for both the standard and delayed-maturers. Only the accelerated-maturers at 14.0 had a variance lower than either CA or ESA. Those results confuse an already puzzling situation. The shuttle run test appears uneffected by biological maturation.

At this point in analyzing the data, I realized that the analysis was not what I wanted. True, I wanted to classify the boys at each CA level on the basis of the relationship between their CA and ESA. However, crossing CA and ESA intervals did not provide that. Recall that the month range for the 13.0 year old is 151 to 163 months. If a boy's CA was 151 months and his ESA was 150 months, then the analysis classified him a delayed-maturers. Obviously, the analysis mis-categorized him. Further, if a boy's CA was 151 months and his ESA was 163 months, then the analysis classified him a standard-maturer . A twelve month difference in CA and ESA should not be a standard-maturer. Therefore, I re-programmed the analysis to subtract ESA from CA. If the difference was -6.00 or more, then I had the analysis classify the boy as an accelerated-maturer. If the difference were between -6.00 and +6.00, then the analysis classified the boy as a standard-maturer. And, if the difference were +6.00 or more, then the analysis classified the boy as a delayed-maturer.

Except that the analysis classified the boys by this more precise procedure, Tables O-9 through O-13 parallel Tables O-4 through O-8.

Table O-9

The sit-ups variances across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	924.62	814.40	814.40	694.72*	500.21*	925.36
Standard	976.76	679.39*	870.26	885.60	480.51*	408.31*
Delayed	------	1115.61	869.84	709.35*	629.13*	887.32

Table O-10

The pull-ups variances across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	8.20*	4.97*	6.79*	12.61	13.11	19.88
Standard	6.44*	5.91*	8.25*	8.99*	15.08	14.52
Delayed	----	7.66*	7.85*	8.39*	17.32	9.18*

Table O-11

The standing long jump variances across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	118.65	78.16*	83.19*	98.70	91.48*	264.97
Standard	35.13*	66.43*	72.36*	113.21	131.87	96.19
Delayed	------	56.14*	68.58*	70.73*	95.68	93.60*

Table O-12

The vertical jump variances across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	14.16	6.89*	8.02*	10.84	10.02	12.73
Standard	3.51*	5.16*	7.25*	10.90	8.81	7.62*
Delayed	-----	6.39*	5.35*	6.37*	10.22	10.80

Table O-13

The shuttle run variances across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	1.87	1.32	0.84*	0.73*	0.75*	0.76*
Standard	0.59*	0.67*	0.90	0.94	0.74*	0.73*
Delayed	----	1.00	0.88	0.74*	0.64*	o.89

A comparison of the tables revealed the following:

(1) For the sit-ups test, the more precise procedure found more homogeneous classifications for the accelerated-maturers at 13.0, 14.0, and 15.0, the standard-maturers at 12.0, and the delayed-maturers at 13.0 and 15.0. To the list of variances that were lower than either CA or ESA, the new procedure added the standard-maturers at 12.0.

(2) For the pull-ups test, the more precise procedure found more homogeneous classifications for the accelerated-maturers at 12,0, 13.0, 14.0, and 15.0, the standard-maturers at 13.0, 14.0, and 15.0, and the delayed-maturers at 14.0. The new procedure had the same groups in its list of variances that were lower than either CA or ESA. Additionally, compared to the CA variances the reduction in variance of the accelerated-maturers at 12.0 was 50%.

(3) For the standing long jump test, the more precise procedure found more homogeneous classifications for the accelerated-maturers at

at 12.0, 13.0, 14.0, and 15.0, the standard-maturers at 15.0, and the delayed-maturers at 12.0. To the list of variances that were lower than either CA or ESA, the new procedure added the accelerated-maturers at 15.0, but lost the delayed-maturers at 15.0. Although the variances for all of the accelerated-maturers groups decreased, their relationships to the standard and delayed-maturers did not change.

(4) For the vertical jump test, the more precise procedure found more homogeneous classifications for the accelerated-maturers at 12.0, 13.0, 14.0, and 15.0, the standard-maturers at 14.0, and the delayed-maturers at 14.0 and 15.0. To the list of variances that were lower than either CA or ESA, the new procedure added the accelerated-maturers at 13.0, but lost the standard-maturers at 15.0.

(5) For the shuttle run test, the more precise procedure found more homogeneous classifications for the accelerated-maturers at 13.0, 14.0, and 15.0, the standard-maturers at 12.0, and the delayed-maturers at 13.0, 14.0, and 15.0. To the list of variances that were either lower than either CA or ESA, the new procedure added the accelerated-maturers at 13.0 and the delayed-maturers at 12.0.

Before examining the mean values for the new classifications, we should examine the distribution of the 1578 subjects (Table O-14).

Table O-14

The distribution of the 1578* subjects across the six CA intervals by the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	11	37	166	178	72	12
Standard	17	96	231	170	108	30
Delayed	0	47	156	148	67	16

*Sixteen subjects were in the 17.0 CA interval

If physical education teachers construct their classes to meet the needs of the standard-maturers, then they will not meet the needs of 58% of the boys at the 13.0 CA level.

Tables O-15 through O-19 provide the mean values for the five motor proficiency tests across the CA intervals.

Table O-15

The sit-ups mean values across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	41.27	54.35	69.71	76.85	85.11	65.58
Standard	55.53	50.98	65.19	72.04	86.46	84.97
Delayed	-----	54.32	64.46	77.42	84.51	82.88

Table O-16

The pull-ups mean values across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	2.00	1.76	2.92	4.16	6.36	4.33
Standard	2.24	2.59	2.95	3.51	5.92	7.97
Delayed	----	3.34	3.46	3.25	4.99	6.13

Table O-17

The standing long jump mean values across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	55.36	59.81	66.04	71.78	76.31	69.33
Standard	57.59	59.03	63.49	66.72	71.90	76.53
Delayed	-----	59.83	62.13	64.64	69.16	69.56

Table O-18

The vertical jump mean values across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	10.82	12.95	14.30	16.31	17.42	15.00
Standard	11.47	12.42	13.48	15.03	15.94	16.37
Delayed	-----	12.70	12.85	13.78	14.55	14.44

Table O-10

The shuttle run mean values across the six CA intervals for the more precise accelerated, standard, and delayed-maturers classifications

	Chronological Age					
	11.0	12.0	13.0	14.0	15.0	16.0
Accelerated	11.45	11.51	10.81	10.39	10.46	10.91
Standard	11.55	11.33	10.97	10.70	10.55	10.34
Delayed	-----	11.31	11.04	10.88	10.76	10.96

For the sit-ups test, while within each category the mean values increase across CA, within each CA interval the pattern varies. The accelerated-maturers have the highest mean values at 12.0 and 13.0. At 14.0, the delayed-maturers have a slight advantage over the accelerated-maturers for the highest mean value. And, the standard-maturers have a 1.3 sit-up advantage over the accelerated-maturers who lead the delayed-maturers by .6 situps at 15.0. Apparently, biological maturation does not effect the sit-ups test.

For the pull-ups test, the results display a paradoxical pattern. From 12.0 to 13.0, one pattern appears and from 14.0 to 15.0 the opposite pattern appears. Between the 13.0 and 14.0 CA intervals, the pattern reversed. At 12.0 and 13.0, the delayed-maturers have the highest mean values and the accelerated-maturers the lowest. However, at 14.0 and 15.0, the accelerated-maturers have the highest mean values and the delayed-maturers the lowest. The accelerated and standard-maturers both show mean value increases from 12.0 through 15.0. But, the mean values for the delayed-maturers are essentially equal at the 12.0, 13.0, and 14.0 CA levels followed by an increase at 15.0.

For the standing long jump test, within each category the mean values increased across CA. However, the mean values for the three categories were essentially equal at the 12.0 level. Thereafter, the accelerated-maturers displayed a clear advantage. Biological maturation definitely effected the results of the standing long jump test.

For the vertical jump test, the pattern followed the standing long jump results. And, the shuttle run test followed the standing long jump and vertical jump results. Whereas the shuttle run variances defied interpretation, its mean values demonstrated that biological maturation effects the shuttle run test.