A B S T R A C T   O F   T H E S I S

DEVELOPMENT

OF A

PERFORMANCE TEST

FOR

SENIOR HIGH SCHOOL MACHINE SHOP

Submitted by

Richard B. Smith

Colorado

Agricultural and Mechanical College

Fort Collins, Colorado

August,

DEVELOPMENT OF A PERFORMANCE TEST
FOR
SENIOR HIGH SCHOOL MACHINE SHOP

There are a number of elements in industrial arts teaching that are measurable. Manipulative skill is one of these elements. All industrial arts teachers undertake to develop manipulative skill, but in varying degrees. To measure the effectiveness of teaching in this respect we must have reliable tools of evaluation. At the industrial arts level practically nothing has been done to develop evaluation tools for manipulative skill in machine shop teaching. The problem of this study seems to be one of the first steps in the development of proper evaluation tools in this area of the industrial arts program.

Problem

How can a performance test for use in senior high school machine shop be developed?

Problem analysis.--In working out this problem these questions must be answered:

1. What has been done in the field of performance testing?

2. What operations should be included in a performance test for pupils with from 200 to 400 periods (150-300 hours) of machine shop instruction?

3. What controls must be exercised in administering the test?

4. What proceduresmust be used to determine the validity of the test?

5. What procedures must be used to determine the reliability of the test?

Delimitation.--The study will be carried on in the senior high schools of Connecticut that offer 400-500 periods of machine shop instruction to pupils in the last three years of high school. This group will include about ten high schools.

Definition.--By performance test is meant a test in which the pupil uses the tools and machines of the machine shop to alter the size, shape, and surface quality of metal.

Throughout this study it must be borne in mind that any of the suggested procedures that grow out of it must be simple enough to seem worth while to a high school industrial arts teacher. Time-consuming statistical methods will not be a practical answer to any of the questions. It should further be noted that the problem and problem analysis have been written so that this is a study of the method of development of a performance test. The actual test used is a vehicle for the study rather than an end in itself.

Methods

Eleven senior high school machine shop teachers cooperated with the writer in the development, administration and criticism of the test.

Selection of operations.--First an inclusive check list of machine shop operations was prepared from

1. Machine Tool Operation, 2v. by Henry D. Burghardt.

2. Machine Shop Training Course, 2v. by Franklin D. Jones.

3. Improving Instruction in Industrial Arts, a bulletin published by the American Vocational Association.

This check list was then reduced to include only those machine shop areas in which the cooperating teachers gave instruction. The teachers next selected from this check list 21 operations that they thought might be combined to make up a test problem.

Preparation of the test.--The test manual which was prepared included the following items:

1. Blueprint of the performance test problem

2. Individual record form

3. Instructions to test administrators on preparation for the test

4. Equipment for the test

5. Instructions to test administrators on administering the test

6. Instructions to pupils

7. Instructions to test administrators on scoring and use of the individual record form

8. A blueprint of test gages

The blueprint, equipment list and instructions to pupils were submitted to the cooperating teachers for criticism and correction.

Administering the test.--The test was administered in 1948 to 97 pupils and in 1949 to 75 pupils. These pupils were from six high schools. The data for the study of the test were taken from the individual records of 95 pupils, 20 of whom were tested in 1948 and 75 in 1949. Thirty one cases were retested in 1949. Thirty test pieces were reinspected twice to check the consistency of inspection. The teachers who administered the test in 1949 submitted criticisms of the inspection procedure.

Master data sheets.--The data that were gathered through the administration of the test and the multiple inspection and criticism of the inspection methods were transferred to three master data sheets.

Analysis of data

The data were analyzed in relation to item validity, content validity and reliability.

Item validity.--Indices of item validity for each item on the individual record form were obtained by the biserial r, by a method published by E. L. Clark in 1928 and by a method published by Lindquist and Cook in 1933. Clark's formula was

$$I.V. = \frac{P - D}{1.00 - D}$$

I.V. is the index of validity

D is the per cent of the whole group failing to respond satisfactorily

P is the per cent of the criterion group failing to respond satisfactorily (The criterion group is the bottom D percent of the whole group)

In Lindquist and Cook's method the I.V. was the ratio between the number of satisfactory responses in the upper and lower quarters of the group. The rank-difference correlation between the biserial r's and Clark's I.V.'s was 0.772 and that between the biserial r's and Lindquist and Cook's I.V.'s was 0.880.

The biserial r's were then compared with the ratios for inspection consistency and the teachers' ratings of the items for inspection procedure. The inspection consistency ratios were obtained by a double reinspection of 30 cases. The number of entirely consistent inspections was divided by 30 to obtain the ratio for each item. The base for the teachers' rating was 8. One item received a teachers' rating of 0. The rank of this item among 32 items was 28 by biserial r and 30 by inspection consistency ratio. There were three items that received a teachers' rating of 4. Their biserial r ranks were 29,26 and 3, and their inspection consistency ranks were 13,27.5 and 27.5.

<u>Test validity</u>.--The first approach to test validity was through a study of the averages and standard deviations of test scores of pupils in different school

groups. The designations of the groups and the number of cases in each group are as follows: B-200 39 cases, B-400 17 cases, D-200 20 cases, D-400 13 cases and E-200 6 cases. The letters B, D, and E represent different schools. The numbers, 200 and 400, indicate the approximate number of periods of instruction the members of the group had had. The scores that were taken from the individual records were in terms of quality, that is accuracy and finish, and speed. In the comparisons of the scores for quality alone of school groups at the same level of instruction, there was a significant superiority of one group over the other in each of the four pairings. In the similar comparisons between groups at different levels of instruction within the same school, there was not in either case a significant superiority of one group over the other. In the comparisons of the scores for speed alone of school groups at the same level of instruction, there was a significant superiority in three of the four pairings. The same schools were not superior in this comparison as in the first case. In comparing groups within the same schools for speed, there was significant superiority of one group each time, but, in one case, it was the 200 period group, and, in the other case, it was the 400 period group. When the scores for quality and speed were combined and similar comparisons made, in three of the four cases of groups at the same level, there was a significant difference between averages. In the fourth case the difference was nearly significant. In the comparisons between groups within the same school, the 400 period group was barely significant in one case. In the other case the superiority was not significant. In each of the three categories of scores, that is quality alone, speed alone, and quality and speed combined, the scores from one school showed greater variability than the corresponding scores from either of the other schools.

The pupils were ranked by their regular shop grades. They were also ranked by their test scores for quality, for speed, and for quality and speed combined. The rank-difference correlations of grades with each of the test scores were found. All five of the correlations of grades and quality alone were positive. One was significant. Three of the correlations of grades and speed scores were positive, one being significant. Two of these correlations were negative. All of the correlations of grades and combined quality and speed scores were positive. Two of these correlations were significant.

Three shortened forms of the test were then made. The first was made by removing from consideration the test results on the six items having the lowest biserial r's. The second form was made by removing the six lowest items according to Clark's I.V.'s and the third by removing the six lowest items according to Lindquist and Cook's I.V.'s. The scores for quality alone on the shortened forms of the test were then correlated with the pupils' grades. In no case did a shortened form of the test have a significantly higher correlation than the original form.

In making up the original scores for the test, the weight of the score for speed had been calculated by a formula developed by R. S. Hunter in 1945. Modified to fit the number of items in this test, the formula is

$$T.S. = \frac{(\text{Mean time} - \text{Working time}) \times 32}{3 \times \text{Mean time}}$$

The time score may be either positive or negative, depending on the relative value of the 'working time'. To check the validity of the weight for time in this formula, new time scores were calculated using 4, 2, and 1 as constants instead of 3. New rank-difference correlations of grades and combined scores for quality and speed were made. Two of the correlations obtained from the use of 2 instead of 3 as a constant were superior; two were inferior, in neither case significantly. As for the correlations obtained with the constants 4 and 1, three out of four were inferior in each case. The differences were not significant.

Reliability.--Reliability coefficients for the test were calculated by four methods for the 95 cases and by five methods for the 31 retest cases. The methods were (1) retest, (2) rational equivalence, (3) odd-even split half, (4) equivalent halves on the basis of expert opinion, and (5) equivalent halves on the basis of item difficulties. The range of the r's for 95 cases was from 0.824 to 0.894. The range for 31 cases was from 0.795 to 0.903. None of the differences between the r's was significant.

Discussion

Approximately a dozen performance tests have been published. Two agencies concerned with teacher training and certification have used performance tests to classify prospective teachers.

Operations for the test.--Six criteria of a good operation for a performance test were identified in this study. They are:

1. The operation must be representative of the work done in the class.
2. The operation must be one that can be performed by anyone in the class.
3. The operation must discriminate between pupils of different total achievement.
4. The operation must be one that can be performed with simple equipment.
5. The operation must be easy to administer.
6. The operation must be easy to check or rate by mere measurement.

The selection of operations, in the first instance, must be based on expert opinion. In the present study, the operations were selected cooperatively by nine machine shop instructors from a check list that was based on three well known books.

After an operation has been made part of a performance test, its validity as an item in the test must be critically studied. The simple method suggested by Lindquist and Cook gave item validity ranks that had a correlation of 0.880 with the ranks by biserial r. Since Lindquist and Cook's method involves but a fraction of the time needed to calculate biserial r's for each item and has a significant correlation with the biserial r, the indication is that it may safely be used by industrial arts teachers in investigating the validity of the items in a performance test.

Controls.--The conditions under which a performance test is administered must be rigidly standardized. Standardization should cover the number, condition and arrangement of tools and machines, the instructions to give to the pupil, and the methods that are to be followed in inspecting the finished test pieces. The methods of inspection set up for this test were critically studied in relation to item validity. A coincidence of low rating by teachers of the inspection methods, poor consistency in actual inspection and low item validity indicates that inadequate inspection methods may be a contributing factor to low item validity.

Test validity.--The differences between the averages of groups of pupils in different schools were generally significant. The differences between groups of pupils at different levels of instruction within the same school were generally non-significant. This indicates that the test measures some factor in the instruction that these pupils have received which varies considerably between schools. The lack of significance between groups within the same school suggests that there is some factor other than manipulative skill that distinguishes pupils in the advanced group from those in the elementary group.

The correlations between pupils' regular shop grades and the quality scores on the test were all positive, the correlations of grades with speed scores were inconsistent, some were positive; some were negative. The combined scores for quality and speed on the test had two significant correlations with the pupils' grades. All five of the correlations in this group were positive. These correlations indicate that speed alone is not a valid score for a performance test. They also indicate that the most valid results from this test are obtained by combining quality and speed scores. The weight given to speed by Hunter's formula, which was used in this study, seems as valid as any tested.

Reliability.— The fact that the coefficients of reliability which were obtained by employing five different methods, had no significant differences between them indicates that any one of the five methods may be used to obtain the reliability coefficient of a performance test.

Suggestions for further study.-- No extensive investigation has been carried on as to the relative merit of the use of fixed tolerances or sliding scales in rating performance tests. It may be that for school instructional purposes the sliding scale would prove superior to the system of fixed tolerances as used in this study.

What are the diagnostic values of the performance test for the teacher? How can it be used to reveal weak spots in the instruction being offered in any industrial arts shop?

There are interesting possibilities for study in the relationship between the performance test and intelligence, mechanical aptitude and post-school success in fields related to the performance test.

T H E S I S

DEVELOPMENT

OF A

PERFORMANCE TEST

FOR

SENIOR HIGH SCHOOL MACHINE SHOP

Submitted by

Richard B. Smith

In partial fulfillment of the requirements

for the Degree of Master of Education

Colorado

Agricultural and Mechanical College

Fort Collins, Colorado

August, 1949

COLORADO AGRICULTURAL AND MECHANICAL COLLEGE

August 1949

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY RICHARD B. SMITH

ENTITLED DEVELOPMENT OF A PERFORMANCE TEST FOR SENIOR HIGH SCHOOL MACHINE SHOP

BE ACCEPTED AS FULFILLING THIS PART OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF EDUCATION

MAJORING IN INDUSTRIAL ARTS EDUCATION

CREDITS 6

Harvey D. Moffle
In Charge of Thesis

APPROVED George H. Henney
Head of Department

Examination Satisfactory

Committee on Final Examination

T. H. Evans — Harvey D. Moffle
W. R. McClanahan — George H. Henney
S. L. Coover

David H. Morgan
Dean of the Graduate School

ACKNOWLEDGMENT

The writer wishes to express his appreciation to the following members and former members of the faculty of Colorado Agricultural and Mechanical College, Fort Collins, Colorado, for their assistance in the preparation of this manuscript:

Prof. George F. Henry, Head of the Department of Industrial Arts and Mr. Harvey D. Waffle, Advisor in Research, for their suggestions and criticism of the manuscript.

Prof. Andrew G. Clark, Head of the Department of Mathematics for his suggestions and criticism of the statistical work.

Dr. James G. Hodgson, Librarian, for his help in the preparation of the bibliography.

Dr. George A. Comstock for his assistance and advice in the preparation of the problem analysis.

The writer also wishes to acknowledge the assistance given him by industrial arts teachers of machine shop in Connecticut, especially Byron A. Berry and Edward D. Busby of Bristol High School and Theodore O. Kennedy of Norwich Free Academy, Norwich, Connecticut.

TABLE OF CONTENTS

TABLE OF CONTENTS.--Continued

LIST OF TABLES

Chapter I

INTRODUCTION

Manipulative skill in the handling of tools and machines is one of several measurable factors in industrial arts. Some of these factors are ability to plan, and ability to read technical symbols, mastery of occupational and technical information, and personality traits such as cooperativeness and readiness to assume responsibility. Teachers of industrial arts have assumed the development of these factors as a part of their responsibility as teachers. Statements of objectives of industrial arts over a period of two decades have included the development of "a certain amount of skill" along with others of these measurable factors.

Generally, the estimate of the degree of skill developed has been based on the judgment of the individual teacher. This has often been influenced by subjective factors. According to Newkirk and Greene (16), it has varied from shop to shop and from teacher to teacher. At best, this judgment has been based on objective measurement or rating of the finished product of the pupil's work. In most cases who can say with any degree of certainty what features of the finished product are the result of the pupil's own effort and ability and what features are the result of the suggestion, advice, specific stimulation or even actual work of the teacher? So, even with perfect evaluation of the finished product we still may not have a true estimate of the ability or skill of the individual pupil. To obtain such a true estimate it would seem that a test problem, which all pupils would make under uniform conditions as to equipment and help from the teacher, is essential.

Such a test problem, provided it be valid and reliable, would serve several purposes for the industrial arts teacher. In the first place it might serve as a check on the regular grading system used by the teacher. It might suggest a change in emphasis on speed, accuracy, technique, planning, or personality factors such as initiative, resourcefulness and perseverance. Then the test might be diagnostic in relation to the instruction that the teacher had been offering. It might reveal weaknesses of which the teacher had not been aware. The results obtained could be used as one factor in predicting probable future success in the industrial area represented by the performance test.

If such a performance test should be used by instructors in several shops under uniform conditions as to previous instruction, equipment and administration of the test, it would provide each teacher with a yardstick for comparing the educational product of his classes with standards set by the group as a whole.

The performance test might be used as a criterion for the evaluation of other tests. It could be used to establish the validity of various aptitude and intelligence tests that are used in the selection of pupils to enter technical areas of instruction. It could be used in the development of paper-and-pencil tests designed for the measurement of achievement in different shop areas.

The first step towards the successful use of performance tests in industrial arts teaching is to explore the methods needed to construct a good test in senior high school machine shop, that is, to seek an answer to some questions about the development of a performance test. The actual machine shop performance test developed and used is a vehicle for the study rather than an end in itself.

Throughout the study it must be borne in mind that any of the suggested procedures that grow out of this study must be simple enough to seem worth while to a high school industrial arts teacher. Time-consuming statistical methods will not be a practical answer to any of the questions.

Problem

How can a performance test for use in senior high school machine shop be developed?

Problem analysis.--In working out this problem these questions must be answered:

1. What has been done in the field of performance testing?

2. What operations should be included in a performance test for pupils with from 200 to 400 periods (150-300 hours) of machine shop instruction?

3. What controls must be exercised in administering the test?

4. What procedures must be used to determine the validity of the test?

5. What procedures must be used to determine the reliability of the test?

Delimitation.--The study will be carried on in the senior high schools of Connecticut that offer 400-500 periods of machine shop instruction to pupils in the last three years of high school. This group will include about ten high schools.

Definition.--By performance test is meant a test in which the pupil uses the tools and machines of the machine shop to alter the size, shape and surface quality of metal.

Chapter II

REVIEW OF LITERATURE

The material that has been published up to the present time deals largely with performance tests in a general way and contributes little toward the answers to the specific questions of validity and reliability. They present, rather, generalizations relative to the need, desirability and feasibility of performance tests. They also give some general principles to follow in the construction, administration and scoring of performance tests.

Desirability and feasibility of performance tests

Chapman (5), 1921, said that trade performance tests are necessary because of the inaccurate results from interviews. The performance test is superior when manual dexterity is a larger part of trade proficiency than knowledge of terms. The scope of the performance test is unlimited. It is superior to a tryout given by a foreman since it is not a chance selection of operations. It has standards for administration and for judging. The standards are the same from month to month. They are independent of the examiner. On the other hand, a performance test cannot cover many operations and therefore may be unreliable.

This study by Chapman was directed toward the use of the performance test in industry, but it has pertinence to the present study in that it points out many of the details in the development of a performance test and its administration.

Toops (22) said, in 1921, that the decision as to whether a pupil was ready for promotion has been based on a subjective estimate of the teacher aided by performance on the job. This decision, however, could not be duplicated by another person who did not know the pupil personally.

He further said that estimates of a pupil's readiness to advance to a job in industry or to an advanced school are made by vocational schools. When based on trade proficiency, these estimates would be improved by more accurate methods of measuring trade skill, such as performance tests. However,

> In most trades,....performance tests have yielded _per unit of time spent in examination_ or original assembly and standardization but a fraction of the examination value yielded by oral or picture tests. (22:30)

The University of Pittsburgh, Vocational Teacher Training Staff (19), in 1934, said in a study of their program of performance tests:

> We do claim two distinct values for the tests. First, they are definitely diagnostic; second, they are prognostic within certain limitations. Their forecasting efficiency cannot be known with the present data. The yardstick may be crude, but its measure is far more satisfactory than a pure guess. (19:30)

Newkirk and Greene (16), 1935, said that the reliability of many early manipulative tests was low and efforts to measure manipulative skill had not been as successful as the measurement of information by the use of the objective pencil-and-paper tests.

In discussing a study they had made of projects in woodworking and drawing, they said that the results indicated the independent ratings made by teachers were highly unreliable and introduced "serious errors in measurement". (16:7)

> It does not seem likely that a scientific method of instruction can be developed in any field without suitable measures of achievement and abilities. (16:29)

Stead and Shartle (17), 1940, said that trade performance tests could be developed in many occupations in addition to typing and stenography. Experiments were being carried on with various types of business and calculating machines.

Greene, Jorgensen and Gerberich (9), 1943, said that the development of standardized tests to rate the proficiency of pupils in trade preparatory training seemed entirely practical, although very little progress had been made.

These writers who have discussed performance tests seem to consider them desirable instruments to be put in the hands of teachers, although Toops (22) does raise the objection of the time involved in construction and administration of this type of test as contrasted with paper-and-pencil and picture tests.

Performance tests already developed

The number of performance tests that have found their way into print up to the present time seems very limited when one considers the general approval given to this type of test.

Bawden (2), in a circular published in 1919, quoted Lt. Col. W. V. Bingham as saying that the army had developed performance tests for about thirty mechanical trades and occupations.

Chapman (5), 1921, published ten of the army performance tests with the permission of the Adjutant General. They were for:

1. Wood pattern maker
2. Pipe fitter -- Steam fitter
3. Sheet metal worker -- General
4. Electrician -- Interior wireman
5. Machinist and mechanic -- Lathe operator
6. Machinist and mechanic -- General
7. Blacksmith -- General
8. Stenographer and typist -- Typist
9. Stenographer and typist -- Stenographer
10. Chauffer -- Truck driver

Number five in the list above was a test in measurement and number six a performance test in bench work. Each test consisted of four parts:

1. Test equipment list
2. Instructions to the examiner
3. Instructions to the candidates, including a blueprint
4. Instructions to the scorer, including a basis for scoring, and a rating scale for placing candidates in the categories of novice, apprentice, journeyman or expert

Newkirk and Greene (16), 1935, published several simple performance tests including one in beginning woodworking. This test had two forms. The correlation of the two forms in thirty cases was "as high as 0.90". (16:148)

Dr. E. J. Gannon, of the Board of Examiners of the Board of Education of the City of New York, in a personal letter to the writer 1/ submitted a list of 22 trades that were the subjects for examination. He also included a performance test in machine shop work dated September 27, 1941. This test consisted of:

1. Instructions to the candidate
2. A mimeographed sketch of the test problem
3. A rating sheet

Hunter (11), 1942, published a short performance test involving the following bench operations in machine shop work: layout, sawing, filing and measurement with

---

1/ Appendix B

scale, micrometer and protractor.

Hunter (12), in 1945, when making a study of certain mechanical aptitude tests for guidance in selecting pupils for vocational machine shop classes, used performance tests in machine shop as a criterion for rating the predictive value of the other tests. The test for sophomores consisted of two parts, both of which involved bench operations. The junior test included bench and lathe work with thread cutting.

Holsinger (10), 1947, published a performance test in basic electricity. The test material included a list of tools and materials needed. The test also had detailed directions for the pupils to follow. It was "a test on accuracy and workmanship". (10:206)

It should be noted that the tests published by Chapman (5) were designed for use with adults in order to classify them as to their advancement in a trade. The tests used by the Board of Examiners, Board of Education, of the City of New York, are for the purpose of rating candidates for teaching positions. The others are performance tests for pupils of junior and senior high school age.

Selection of operations

The records of past investigations in performance testing offer only general principles relative to the selection of operations to be included in a performance test in senior high school machine shop.

Chapman (5), 1921, said that two requisites of a performance test were: first, the task should be as typical as possible of the actual work in the trade, and second, the test must be such that it can be used for anyone in the trade without regard to his development in the trade.

> An operation is of no service as the basis of a performance test unless (1) it will differentiate between different levels of trade ability; (2) it can be administered with fairly simple equipment; (3) it can be administered and be rated in an objective manner by a non-tradesman. (5:273)

A performance test is of most significance when it involves the largest number of significant operations. The length of a performance test is a compromise between the desire for reliability and the practical limits imposed by shop conditions.

Toops (22), 1921, said that performance tests:

> ....have the merit not possessed by ordinary shop jobs, of being so chosen as to require a maximum amount of manipulation of tools, use of trade knowledge and trade skill. (22:7)

and

> ....such tasks must be chosen as will lend themselves to easily and objectively measurable scoring points....The best performance test is one in which mere measurements of the product will yield the desired differentiation.... (22:28-9)

Hunter (12), 1945, found that a longer test or a combination of two tests gave more valid results, increasing the correlation with shop grades from 0.64 to 0.71.

Controls

Controls are those instructions, given in detail to the test administrator, to assure uniform conditions at different times and places.

Chapman (5) said, in 1921, that the conditions under which a performance test is given should be controlled by having definite instructions provided for equipment, the examiner, the candidate and for scoring. The equipment, tools and materials must be reduced to the smallest possible quantity. They must be capable of standardization. Standardization should cover the number of tools, their arrangement and their condition. All instructions that are to be given to the candidate should be written out. The examiner should not be allowed to deviate in the least from the printed instructions. Care must be exercised to see that the candidate does not receive outside help. Objectivity in scoring should be obtained by preparing standardized instructions for scoring the test.

Bingham (3), 1942, said that a testing program would yield disappointing results unless the tests were well administered. The physical setting for the test should be carefully prepared without distracting influences. The subject (pupil) should be psychologically prepared for the test. The examiner should know the directions well enough to be able to give them without hesitation. "The first rule in test administration is to secure and to maintain the conditions prescribed for the test." (3:238)

Controls for scoring.--Toops (22), in 1921, said that the use of a series of limit gages was superior to the use of micrometers with the possibility of variations between scores.

> Statistically, the ordinary limit gages of a dimension ±.001 are as poor a measure of a product for test purposes as dividing all humanity into short, tall and medium people. Thus partial credits are allowable, logical and necessary for best results in performance tests....(22:28-9)

The University of Pittsburgh, Vocational Teacher Training Staff (19), 1934, in discussing questions relating to scoring, said that they gave consideration to time and skill in constructing performance tests. "The question of 'tolerance' was carefully considered." (19:23) Some trade committees used the industrial practice of a specific tolerance within which a product was accepted and above or below which a product was rejected. Other committees used a sliding scale which they termed "instructional practice". The machine shop committee rated candidates by both scales.

> It is interesting to note that in using either scale there was no difference in classification in eight of the nine candidates. (19:27)

The discussion on controls for scoring shows that there is still disagreement on whether a sliding scale of tolerances or the standard industrial practice of tolerances should be used in scoring the results of performance tests.

Validity

A considerable part of the literature on validity is related to tests in general rather than to performance tests specifically. Content validity is one aspect of test validity.

Orleans, 1937, said: "In the last analysis the sources of validity of an achievement test....are the opinions of individuals." (18:46)

The University of Pittsburgh, Vocational Teacher Training Staff (19), in 1934, said that there were naturally questions as to the validity and reliability of the performance tests they had used. They had made no statistical studies because of insufficient data, but identified at least five significant criteria for evaluation and interpretation.

> First. ....expert judgment....In the case of the trade tests, three experienced trade teachers adjudged as experts in their fields sat on committees. These committees prepared....performance tests.
>
> Second. Evident correlation of performance test with written test and an agreement between actual performance and the experience record.
>
> Third. Written and performance tests proved themselves to be diagnostic.
>
> Fourth. Classification of the candidate by his employer in industry.
>
> Fifth. Consistency of several rating scales based on the same performance test. (19:29-30)

Newkirk and Greene (16), 1935, said that validity is often determined by the extent to which it calls into play skills and abilities which experienced observers consider fundamental. Some of the "skills and abilities" that might have reference to a machine shop performance test are accuracy, quality of finish, technique, speed of performance and planning ability.

Speed.--Chapman (5), 1921, said that a performance test should be scored for both quality and speed.

Toops (22), 1921, also maintained that a performance subject should be rated on the quantity and quality of the work done.

> The time required for the total operation must always be carefully taken and preserved, quality and quantity of production being the two generalized measures of the workman's skill in fashioning his product. (22:29)

Newkirk and Greene (16) said, in 1935, that there might be four kinds of performance test exercises: (1) tests of quality or accuracy, (2) tests of identification of tools and materials, (3) tests of technique in handling tools and (4) tests of speed. Their discussion suggests that these various types should be separate tests. However, in discussing the fourth type they said:

> Speed and accuracy are each variable factors in achievement and performance. Test exercises designed to measure speed or rate of response must present a well defined activity with appropriate standards. (16:122)

Hunter (11), in 1942, gave the following formula for computing the weight of the time score:

$$E = P + (P \times S_q)$$

E is the overall efficiency of the pupil.

P is the total score of the pupil for accuracy and finish.

S_q is the speed quotient and is found from this formula:

$$S_q = \frac{\text{Mean time - Working time of pupil}}{\text{Mean time}}$$

(S_q may be either positive or negative depending on the relative size of the "working time" of the pupil.)

The mean time is for the entire sample.

Columns 6 and 7 of Hunter's table showed the ranking of pupils with and without the correction of the scores for speed and indicate a considerable variation.

Hunter (12), 1945, modified his formula to this form:

$$T.S. = \frac{(\text{Mean time - Working time}) \times 100}{3 \times \text{Mean time}}$$

T.S. is the time score for the individual pupil. It may be either positive or negative and is combined with the pupil's score for accuracy and finish.

In a personal letter to the writer 2/ Hunter

2/ Appendix C

said that his change in the weight of the time score was made to decrease the effect of speed and not to handicap unduly the slow but accurate pupil.

There is general agreement that the speed of performance should be taken into consideration in scoring a performance test. There is not, however, any agreement as to the appropriate weight to give to the time score on any test.

Finish.--Hunter (11), in 1942, showed in his table that finish had been considered in determining the proficiency score for each pupil.

In 1945, Hunter (12) counted finish as ten per cent of the proficiency score, giving the other ninety per cent for accuracy.

Item validity.--The discussions of item validity that follow were all presented with the more conventional types of objective in mind. They are presented here since a method of checking various items of the test for validity that would be satisfactory to shop teachers should be sought in this study.

Clark (6), 1928, presented a formula for the calculation of indices of validity:

$$I.V. = \frac{P - D}{1.00 - D}$$

in which

I.V. is the index of validity.

D is the per cent of the whole group failing to respond satisfactorily.

P is the per cent of the criterion group failing to respond satisfactorily. (The criterion group is the bottom D per cent of the whole group.)

Lindquist and Cook (14), 1933, cautioned against assuming a relationship between the difficulty and the discrimination of an item. They offered several suggestions for determining the discrimination value of an item. The simplest was the following:

I.V. is the ratio of the number of satisfactory responses in the upper and lower quarters of the group.

Wood, Lindquist and Anderson, 1933, said:

> ....the validity of any single item in the test also must depend (within limits) upon the degree to which that item of itself discriminates between pupils of inferior and superior total achievement. (23:16)

Smith (21), 1934, found that items with a very low "biserial r" have an adverse effect on a test but that items with medium or high "r's" are about equal in value in making a test. The lowest 20 per cent might be eliminated this way.

Davis (7) said, in 1946, that indices of discrimination are "only one fallible guide" (7:387) to be used in the selection of test items.

Garrett (8), 1947, said that the "biserial r" is the standard procedure for determining whether the item discriminates between pupils differing sharply in total achievement.

Three of the souces quoted above mention the "biserial r" as one statistical method of obtaining an estimate of the validity of a given item. The other formulas submitted entail much less calculation and may, if they give comparable results, be more practical for the shop instructor who is not primarily interested in research work.

Reliability

There are numerous methods of obtaining an estimate of the reliability of a test. However, there is no indication whether they can be applied safely to a performance test where the items are not arranged in a random order on the test record sheets. In the matter of reliability as well as validity a method which is not too cumbersome must be sought.

Newkirk and Greene (16), 1935, said: "The reliability of a test may be thought of as the consistency with which it performs." (16:134) This consistency is a function of the adequacy of the sampling and of varitions of human responses. The first of these can be controlled in part by exercising care in the selection of items. The items should be selected extensively from the field that the test is intended to measure.

Richardson and Kuder (20), in 1939, gave several formulas, which they took from an earlier article, for use in estimating reliability coefficients by rational equivalence. Two of their formulas follow:

$$r_{tt} = \frac{n}{n-1} \cdot \frac{\sigma_t^2 - \Sigma pq}{\sigma_t^2} \qquad \text{No. 20}$$

$$r_{tt} = \frac{n}{n-1} \cdot \frac{\sigma_t^2 - n\bar{p}\bar{q}}{\sigma_t^2} \qquad \text{No. 21}$$

In the formulas the symbols have the following meanings:

r_{tt} is the r for the test t.

σ_t is the standard deviation of the test scores.

p is the per cent correct of a test item.

q is the per cent incorrect on a test item.

n is the number of items.

$\bar{p}$ is the average per cent of correct answers given to each item, computed by dividing the mean by the number of items.

Greene, Jorgensen and Gerberich (9), 1943, discussed the "Footrule" coefficient as a simple method of obtaining an estimate of reliability. The formula which they gave is formula No. 21 in the Richardson-Kuder article. (See above)

Garrett (8), 1947, said that there are three procedures in common use for determining the reliability of a test, namely: (1) test-retest method, (2) alternate or parallel-forms method and (3) the split-half method.

The test-retest method is the simplest, but is open to some objections. When the retest closely follows the first testing, transfer, memory and learning may cause correlation to be too high. Making the lapse between test and retest longer will reduce these objections, but added maturity, skill and learning may tend to reduce the reliability coefficient unduly.

The use of the parallel or alternate forms is open to most of the objections to the test-retest method in addition to the difficulty in devising two tests that are equivalent without making them too much alike.

> The split-half method is employed when it is not feasible to construct an alternate form of the test nor wise to repeat the test. This situation occurs with many performance tests. ....The split-half method is generally regarded as the best....(8:382-3)

The method of rational equivalence was offered as an alternative to the above methods, being free from the various objections relating to administration of the tests. Formula No. 20 in the Richardson-Kuder article (See above) was given, as well as this alternative formula:

$$r_{1I} = \frac{n\sigma^2_t - M(n - M)}{\sigma^2_t(n - 1)} \qquad (8:385)$$

Summary

The review of literature answers in whole or in part the first three questions of the problem analysis. It also offers suggestions for the answers to the last two questions.

What has been done in the field of performance testing.--Performance tests have been devised for thirty or more mechanical trades or areas. Performance tests in machine shop have been used by the U. S. Army, by the Board of Examiners of the Board of Education of the City of New York, by the University of Pittsburgh, Vocational Teacher Training Staff and by Hunter in his studies of aptitude tests for machine shop classes. In none of these cases was the test used at the industrial arts level. Holsinger's electricity test and Newkirk and Greene's tests in woodwork were for industrial arts classes.

What operations should be included in the present test.--The operations should be ones that lend themselves to easy, objective measurement. The operations should differentiate between pupils of superior and inferior total achievement. The number of operations should be as great as the practical conditions of shop teaching will allow.

What controls must be exercised.--The controls for the proposed test must cover tools and machines, instructions to pupils and methods of scoring. Opinion is divided as to the use of sliding scales or absolute tolerances in rating performance tests.

Validity.--Content validity is determined by an appeal to the authority of experts or by correlation with some criterion. Dimensional accuracy, finish and speed are considered fundamental elements of machine shop performance. Various methods, including the "biserial r" have been suggested for investigating the validity of individual items.

Reliability.--No one who has used performance tests in machine shop has made any study of methods for obtaining an estimate of reliability. Newkirk and Greene used the equivalent-forms method with a woodworking test. The retest and split-half methods are untried as far as this type of test is concerned.

Chapter III

METHODS

There were three major phases in the data-gathering part of the attack on the present problem. The first phase was that of securing the collaboration of Connecticut industrial arts teachers of machine shop. The second phase -- developing the test -- divided itself into several steps such as the selection of operations for the test, preparation of the test problem and the preparation of detailed instructions for administering and scoring the test. The last phase was the administration of the test to senior high school pupils.

Securing collaboration

Securing the cooperation of other machine shop teachers was extended over a considerable period -- more than half of the two years that were devoted to the study. In the fall of 1947 the teachers listed as industrial arts teachers of machine shop by the State Board of Education, Hartford, Conn. were contacted either personally or by mail 1/.

1/ Appendix E

Twenty teachers were contacted, eleven of whom replied. Eight indicated interest in taking part in the study. One gave a simple negative answer. One reported that his shop was being changed into a general metal shop. One reported verbally that the authorities in his school would not permit the administration of tests for persons outside the school system. During the year (1947-8) one more teacher agreed to administer the test to a limited number of his pupils. During the second year one other teacher joined the study.

Table 1.--PARTICIPATION OF TEACHERS AND PUPILS

Year	Teachers		Pupils	
	Development of test	Administration of test	Starting test	Complete records
1948	9	8	97	66
1949		7	83	75
TOTAL	9	9[1]	180	141

[1]Six teachers administered the test both years.

Developing the test

Selecting operations for the test.--A complete list of machine shop operations was prepared. It was based on three authoritative sources:

1. Machine Tool Operation, 2 vols.
   by Henry D. Burghardt (4)

2. Machine Shop Training Course, 2 vols.
   by Franklin D. Jones (13)

3. Improving Instruction in Industrial Arts
   a bulletin published by the
   American Vocational Association (1)

It was realized that this list would contain operations which were not covered by the majority of the cooperating teachers. The teachers were contacted by mail and asked to list the machines on which they gave instruction. They replied on a form which was typed on a post card 2/. The list of machine shop operations was reduced so that it included only the areas in which instruction was generally given and the list was then mimeographed as a Check List of Machine Shop Operations 3/. This check list with a covering letter 4/ was mailed to the cooperating group of teachers. When the check lists were returned it was found that the teachers had checked far too large a number of operations. The average number was 68.4, the smallest number 18 and the largest number 119.

The writer then arranged an interview with each of the teachers. The purposes of the interview were (1) to arrive quickly at a suitable list of operations for the test, (2) to acquaint the writer with the shops in which

---

2/ Appendix F

3/ Appendix G

4/ Appendix H

the test would be used and , (3) to learn on what basis the teachers had checked the Check List of Machine Shop Operations for the test. A typed form was used to record the results of the interviews 5/. The Check List of Machine Shop Operations was also used. The interviews showed that the teachers had originally checked either all the operations that they taught or all they thought were suitable for inclusion in the proposed test. Only one had checked a limited number of operations that he thought could be combined to make up a single test problem. Another result of the interviews was a list of 21 machine shop operations which could be used as a basis for the proposed performance test problem 6/.

Development of the test problem.--The list of operations mentioned in the last paragraph was used as the basis of a blueprint of an actual test problem. The problem included 15 of the 21 operations as required machining operations and two others as either incidental or optional procedures. One operation, chamfering the threads of the work piece which was not in the list of 21 operations, was included in the preliminary form of the test problem. Machinery's Handbook, 11th edition, was used throughout the study as the authority for standard tolerances where

---

5/ Appendix I

6/ Appendix J

they were applicable. The preliminary form of the blueprint 7/was mailed to the cooperating teachers along with a questionnaire 8/. The preliminary blueprint was also submitted to the teacher of drafting at the Bristol (Conn.) High School. The suggestions and criticisms that were received from these sources were incorporated in the final form of the blueprint of the test 9/.

Development of directions for the test.--Simultaneously with the development of the preliminary blueprint of the test problem, preliminary specifications for equipment and tools needed for the test and instructions to be given to pupils were prepared. The general rules laid down by Chapman (5) were kept in mind during this stage of the work. These two sections of the test manual, which was now beginning to take form, were also submitted to the cooperating teachers for criticism and suggestion. A questionnaire relating to each one was used 10/ 11/.

Instructions for test administration were prepared. These were in two parts, the first relating to preparation for the test and the second to the actual administration of the test.

7/ Appendix K

8/ Appendix L

9/ Appendix D

10/ Appendix M

11/ Appendix N

Material relating to scoring the test was in three parts. The first of these was an individual record form. This was to gather all the data relative to the individual tests: (1) data about the pupils, (2) the time spent on the tests and (3) the record of the inspection of the finished test pieces. The individual record forms were mimeographed 12/. In the preparation of the individual record forms it was decided to ask the teachers to give each pupil's rank within his own class. This was substituted for the pupil's grade as had been originally planned, as the grading systems in the various high schools were not uniform.

The second part of this group of materials was the instructions for the use of the individual record form. They covered the recording of the individual data, the use of the time record, and the inspection of the test pieces.

The third part of this work was the preparation of gages to make the inspection more uniform. All the gages that it was possible to use were not made since some of the tolerances were still tentative. A blueprint showing the gages used was prepared 13/.

All of the above materials were incorporated in a Manual for Experimental Use, a copy of which was sent to

---

12/ Appendix D

13/ Appendix D

each teacher that planned to administer the test during the latter part of the 1947-8 school year. These teachers were also supplied with copies of the blueprint of the test problem and of the individual record forms, as well as a complete set of gages. Copies of these materials appear in the revised 1949 Manual for Experimental Use 14/.

Administration of the test

During May and June, 1948, the Machine Shop Performance Test was first administered to senior high school pupils in several schools in Connecticut. Ten teachers including the writer had either planned or agreed to administer the test in eight different schools. Actually, eight teachers did administer the test to 97 pupils in five high schools. Sixty-six of the individual records were complete enough to be used in preliminary statistical study of the test results. Thirty-eight of these records were of pupils who had completed approximately 200 forty-five-minute periods, or their equivalent, of instruction. The remainder, 28, were of pupils who had completed 400 or more periods of instruction. Six of the eight instructors including the writer filled in a Questionnaire-Criticism of the test and the test procedure 15/. The answers on the Questionnaire-Criticism were summarized and used where

---

14/ Appendix D

15/ Appendix O

possible in making minor modifications in the test manual for use in 1948-9 16/.

During the spring months of 1949 the test was administered to 75 pupils in three high schools. Of these pupils, 57 had had approximately 200 periods of machine shop work in high school and 18 had had approximately 400 periods.

Table 2.--TOTAL NUMBER OF CASES USED FOR STUDY OF TEST

Schools	400 periods of machine shop			200 periods of machine shop			School totals
	1948	1949	Total	1948	1949	Total	
B	12	5	17	8	31	39	56
D		13	13		20	20	33
E					6	6	6
TOTAL	12	18	30	8	57	65	95

The teachers who worked with the test during 1948-9 submitted criticism of the instructions for inspection on a form which was supplied to them 17/.

Retest.--Thirty one of the pupils tested in 1949 were retested. The same procedure was followed as on the original test. The same record of inspection and time

16/ Appendix P

17/ Appendix Q

required for the test was made. The retest was administered about 10 class periods after the completion of the original test.

Reinspection

In May and June, 1949, thirty sample test pieces from three schools were reinspected twice by a teacher in the machine shop of the Bristol (Conn.) High School. This teacher had participated in the administration of the test both years. His reinspections, however, were entirely independent of the original inspections on these test pieces.

Equivalent halves -- Experts lists

In May, 1949, the 32 items of inspection on the individual record forms were divided into groups thought to be equivalent in difficulty to the pupils. The other two teachers in the machine shop of the Bristol (Conn.) High School and the writer made independent pairings of the items. These pairings were analyzed. Six items were found to be paired identically by all three teachers. Twenty items were paired the same by two of the three teachers, leaving six items on which there was no agreement 18/. In conference these last six items were paired. This gave two lists of 16 items each that were thought to be equivalent by these teachers.

18/ Appendix R

Master data sheets

The data gathered through administration of the tests, the reinspection of test pieces and the criticisms of the inspection methods by teachers were transferred to three master data sheets 19/. The data from testing and retesting of pupils and the information about the pupils was entered on sheet number one. The ranks in class of the pupils were written as a ratio, the first member of the ratio was the rank and the second member was the number of pupils in the class. The reinspection data was entered on data sheet number two. The teachers' ratings of the inspection procedure were put on sheet number three.

19/ Appendix A

Chapter IV

ANALYSIS OF DATA

In Chapter II the contributions to this study of previous writings and investigations on tests and performance tests were brought out. Chapter III outlined the methods followed in laying the groundwork for the study and in gathering the data. The present chapter will take up the analysis of the data that have been gathered. The three main concerns will be item validity, content validity, and reliability; the procedures to follow in obtaining estimates of item validity, validity and reliability. The statistical procedures used have followed Garrett (8) unless otherwise noted.

Item validity

Indices of validity were obtained by three methods and the methods compared. The indices of validity obtained by the biserial r method were compared with the ratio for inspection consistency, and with the teachers' rating of the inspection methods.

The biserial r's for the items of inspection on the test range from 0.089 for the 4¼ length to 0.642 for the finish on the 31/32 diameter, Table 3. When Clark's (6) formula was employed, the highest index of validity was 0.791 for the finish on the 31/32 diameter while the lowest was -0.002, again for the $4\frac{1}{4}$ length. According to Lindquist and Cook's (14) method the lowest index of validity was 1.09 for the 31/32 diameter and for the finish on the 12° angle and on the threaded end. The rank-order coefficient of correlation between the ranks of items by their biserial r's and the ranks according to the I.V.'s obtained by Clark's formula was 0.772, while that between the ranks by biserial r's and Lindquist and Cook's method was 0.880. Both of these rho's are significant at the .01 level, as shown in the table in Garrett. (8:299)

Table 3.--VARIOUS METHODS OF FINDING INDICES OF ITEM VALIDITY

Inspection item	r_{bis}		Clark		Lindquist and Cook	
	I.V.	Rank	I.V.	Rank	I.V.	Rank
1	2	3	4	5	6	7
4 1/4	.089	32	-.002	32	1.67	23
3	.320	16	.266	20	2.71	14
1 1/4 - thread	.094	30	.186	22	1.20	30
3/4	.255	18	.326	17	2.00	19
1/4	.097	29	.039	31	1.23	29
31/32	.116	27	.366	15	1.09	32
11/16	.168	24	.176	23	1.63	24
1/16 chamfer	.098	28	.131	27	1.83	20
3/32	.354	11	.285	18	2.83	13
.188	.534	6	.611	10	12.00	5.5
.750	.162	25	.136	26	1.60	25
Thread P.D.	.215	22	.171	24	2.50	17
.075	.219	21	.372	14	2.60	15
1 1/4 hole	.308	17	.705	3	8.00	8.5
3/8 ream	.525	8	.656	7.5	11.00	7
12° angle	.091	31	.329	16	1.18	31
30° chamfer	.569	3	.701	5.5	23.00	4
90° holes	.349	12	.256	21	1.71	21
1/4-20 thread depth	.247	19	.142	25	1.41	27

Table 3.--VARIOUS METHODS OF FINDING INDICES OF ITEM VALIDITY--Continued

Inspection item	r_{bis} I.V.	r_{bis} Rank	Clark I.V.	Clark Rank	Lindquist and Cook I.V.	Lindquist and Cook Rank
1	2	3	4	5	6	7
1/4-20 thread square	.117	26	.082	30	1.43	26
1/64 radius	.243	20	.274	19	1.69	22
Shoulder	.323	15	.110	28	2.50	17
Centering hole	.181	23	.085	29	1.25	28
Centering keyway	.342	14	.432	12	2.50	17
End	.489	9	.656	7.5	5.50	12
Burring	.346	13	.377	13	7.00	10
<u>Finish</u>						
Turning 12° angle	.466	10	.747	2	∞	1.5
Turning 31/32 diameter	.642	1	.791	1	24.00	3
Turning .750 diameter	.583	2	.701	5.5	12.00	5.5
Facing threaded end	.552	4	.620	9	∞	1.5
Facing shoulder	.550	5	.704	4	5.75	11
Facing tapered end	.529	7	.604	11	8.00	8.5

Inspection consistency ratio.--The inspection consistency ratio was calculated from the reinspection data 1/obtained from 30 cases. Each test piece was inspected three times. If all three inspections on an item were checked the same that was counted as a perfect inspection for that item. The cases in which an item was given perfect inspection were counted. This number was divided by 30, the total number of cases reinspected, to obtain the ratio. For example, the first item, 4¼ length, was checked as satisfactory in all three inspections in 14 cases and unsatisfactory in all three inspections in seven cases. This made a total of 21 perfect inspections; 21 divided by 30 gave an inspection consistency ratio of 0.700.

The 1/16 chamfer ranked 28th, Table 4, by biserial r, 30th by inspection consistency and received a composite teachers' rating of 0, on a scale of 8. The 1/4 dimension was 29th by biserial r, 13th by inspection consistency and had a teachers' rating of 4. The ¼-20 thread square ranked 26th by biserial r, 27.5 by inspection consistency and had a teachers' rating of 4. The 30° chamfer also received a teachers' rating of 4 and was 3d in the biserial r ranks and 27.5 in inspection consistency ranks.

1/ Appendix A

Table 4.--COMPARISON OF ITEMS BY BISERIAL r AND INSPECTION CONSISTENCY, AND TEACHERS' RATING OF INSPECTION

Inspection item	r_{bis}		Inspection consistency		[1]Composite rating
	I.V.	Rank	Ratio	Rank	
1	2	3	4	5	6
4 1/4	.089	32	.700	22	6
3	.320	16	.700	22	6
1 1/4 thread	.094	30	.833	9.5	8
3/4	.255	18	.567	29	5
1/4	.097	29	.800	13	4
31/32	.116	27	.933	2	8
11/16	.168	24	.700	22	7
1/16 chamfer	.098	28	.533	30	0
3/32	.354	11	.700	22	6
.188	.534	6	.767	16.5	7
.750	.162	25	.800	13	7
Thread P.D.	.215	22	.933	2	7
.075	.219	21	.867	6.5	5
1 1/4 hole	.308	17	.867	6.5	7
3/8	.525	8	.400	32	7
12° angle	.091	31	.900	4	6
30° chamfer	.569	3	.600	27.5	4
90° holes	.349	12	.867	6.5	7

[1]The highest possible rating is 8.

Table 4.--COMPARISON OF ITEMS BY BISERIAL r AND INSPECTION CONSISTENCY, AND TEACHERS' RATING OF INSPECTION--Continued

Inspection item	r_{bis}		Inspection consistency		Composite rating
	I.V.	Rank	Ratio	Rank	
1	2	3	4	5	6
1/4-20 thread depth	.247	19	.800	13	7
1/4-20 thread square	.117	26	.600	27.5	4
1/64 radius	.243	20	.867	6.5	7
Shoulder	.323	15	.467	31	8
Centering hole	.181	23	.933	2	5
Centering keyway	.342	14	.800	13	6
End	.489	9	.800	13	7
Burring	.346	13	.633	25.5	8
Finish					
Turning 12 angle	.466	10	.833	9.5	8
Turning 31/32 diameter	.642	1	.733	18.5	8
Turning .750 diameter	.583	2	.733	18.5	8
Facing threaded end	.552	4	.768	16.5	8
Facing shoulder	.550	5	.633	25.5	8
Facing tapered end	.529	7	.700	22	8

Test validity

The validity of the test was studied by comparing the averages of different groups of pupils and the standard deviations of the test scores of the different groups. Validity was estimated by correlation of the test scores with the pupils' grades by the rank difference method. The effect of ignoring the inspection record of certain items of low validity was studied as was the effect of changing the weight of the time score.

In the scores for quality alone the D-400 2/ group made the highest average score of 24.15, Table 5, with a SD of 3.25. The lowest group average was 12.95 for the B-200 group with 4.49 as the SD. The average score for all pupils was 17.26 with a SD of 7.71. The test scores of the two groups in school B had larger SD's than those of the other three groups, two in school D and one in school E, with 4.49 and 3.98 as against 2.39, 3.25 and 2.61, respectively.

In the score for time alone, the average for all pupils was 223 minutes. The SD of these scores was 76.7. Again the scores of the two groups in school B had larger SD's, 51.1 and 66.6, than the scores of the other school groups where the SD's were 28.3, 24.9 and 38.2.

2/ In this group symbol the letter D represents one of the schools in which the performance test was administered; the 400 indicates the number of periods of instruction in machine shop that the group has had.

In the combined scores for both quality and speed the average for the whole group was 17.25 and the SD was 7.76. As in the former cases the groups in school B showed greater variability with SD's of 6.12 and 5.57 while the SD's for the other three groups were 2.78, 3.43 and 3.18.

Table 5.--GROUP AVERAGES AND VARIABILITY

School Groups	No. of Cases	Quality		Speed		Quality and Speed	
		M	SD	M	SD	M	SD
1	2	3	4	5	6	7	8
[1] B-200	39	12.95	4.49	207	51.1	13.74	6.12
B-400	17	14.23	3.98	268	66.6	12.05	5.57
D-200	20	23.60	2.39	200	28.3	24.10	2.78
D-400	13	24.15	3.25	160	24.9	27.10	3.43
E-200	6	17.83	2.61	418	38.2	8.60	3.18
All	95	17.26	7.71	223	76.7	17.25	7.76

[1]In the symbol for school groups the letter represents the school and the number represents the approximate number of periods of 45 minutes each of machine shop practice.

When the averages in quality scores of different school groups were compared, Table 6, school D was seen to be superior to school B in both the 200 period group and the 400 period group, where the critical ratios, CR's, were

10.34 and 8.00, respectively. These compared with table entries (8:190-1) of 2.67 and 2.76 at the .01 level of significance. In both cases there was much less than one chance in a hundred that the difference between the groups was due to sampling errors. School D was also superior to school E in the 200 period groups, where the CR of the difference was 4.89 against 2.80 in the table at the .01 level. School E was superior to school B in the 200 period group comparison. In neither school B nor school D was there any significant difference between the 200 period group and the 400 period group. The CR's were only 1.00 and 0.57 respectively. These were much below the table entries for the .05 level of significance, which were 2.01 and 2.04.

In the scores for speed both school B and school D were superior to school E in the 200 period groups. The CR's were 9.38 and 15.14, respectively, which were well above the table entries for the .01 level of significance. There was no significant difference between school B and D in this group since the CR was only 0.56 which was much below the table entry for the .05 level of 2.00. In the 400 period groups, school D was again superior to school B with a critical ratio of 5.48 compared with 2.76 in the table for the .01 level. In the comparison between the different levels of instruction within the same school the 200 period group was superior in school B and the 400 period group in school D. The CR's were 3.67 and 5.77, respectively, as against 2.67 and 2.76 in the table for the .01 level of significance.

In the combined scores for quality and speed, Table 6 -- columns 9 and 10, school D was superior to school B in both groups and to school E in the 200 period group. All the CR's were well above those required for the .01 level of significance. In the intra-school comparisons neither group was superior in school B. In school D the 400 period group was superior at the .05 level of significance.

Table 6.--SIGNIFICANCE OF THE DIFFERENCES BETWEEN GROUP AVERAGES

Groups compared	No. of cases	Significance levels .05	.01	Quality D	Quality t	Speed D	Speed t
1	2	3	4	5	6	7	8
B-200 D-200	39 20	2.00	2.67	[1]D-200 10.65	10.34	D-200 7	0.56
B-200 E-200	39 6	2.02	2.70	E-200 4.88	2.51	B-200 211	9.38
D-200 E-200	20 6	2.06	2.80	D-200 5.77	4.89	D-200 218	15.14
B-400 D-400	17 13	2.05	2.76	D-400 9.92	8.00	D-400 108	5.48
B-200 B-400	39 17	2.01	2.67	B-400 1.28	1.00	B-200 61	3.67
D-200 D-400	20 13	2.04	2.76	D-400 0.55	0.57	D-400 40	5.77

[1]The number of the group showing superiority is placed over the difference in each comparison.

Table 6.--SIGNIFICANCE OF THE DIFFERENCES BETWEEN GROUP Averages--Continued

Groups compared	No. of cases	Significance levels .05	.01	Quality and Speed D	t
1	2	3	4	9	10
B-200 D-200	39 20	2.00	2.67	D-200 10.97	7.67
B-200 E-200	39 6	2.02	2.70	B-200 5.14	1.97
D-200 E-200	20 6	2.06	2.80	D-200 16.11	11.43
B-400 D-400	17 13	2.05	2.76	D-400 15.05	8.27
B-200 B-400	39 17	2.01	2.67	B-200 1.69	0.93
D-200 D-400	20 13	2.04	2.76	D-400 2.39	2.12

The ranks of pupils by their regular shop grades and by the various scores on the test were compared by the rank difference method, Table 7. (8:299,344-7) The rank-order coefficient of correlation (rho) between grades and the combined scores for quality and speed of the B-200 group was 0.63. This was significant at the .01 level. In the comparison of the scores of the B-400 group the rho for grades and speed was 0.79 which was significant at the .01 level. For the D-400 group, the rho's, 0.71 and 0.70, for the correlation of grades with quality scores alone a and with the combined quality and speed scores, respectively, were both significant at the .01 level. None of the other rho's reached significance at the .05 level.

With the B-200 and B-400 groups the rho's for the grades and combined score were considerably higher than the rho's for the comparison of grades and quality alone. With the D-200, D-400 and E-200 groups, on the other hand, the rho's for the comparison of grades with the combined score for quality and speed were lower than that for quality alone, although the drop from 0.71 to 0.70 for the D-400 group was not significant.

Table 7.--RANK-DIFFERENCE CORRELATIONS OF TEST SCORES WITH SHOP GRADES

School Groups	No. of Cases	Quality and Grades	Speed and Grades	Quality--Speed and Grades
1	2	3	4	5
B-200	[1]17	.41	.40	.63
B-400	12	.12	.79	.43
D-200	20	.39	-.18	.17
D-400	13	.71	.24	.70
E-200	6	.33	-.31	.20

[1]The number of cases in some school groups is less in this table than in Table 3 because all of the cases in the earlier table were not ranked by the same teacher the same year. The largest homogenous group was taken as the sample each time.

The six items with the lowest I.V.'s in each of the rankings made in Table 3 were then deleted from the test results 3/. This gave three shortened forms of the test. The first was shortened by the biserial r ranking of the items, the second by the I.V.'s obtained with Clark's formula and the third form by using the I.V.'s obtained with Lindquist and Cook's method. The ranks of the pupils by the quality scores on these shortened forms of the test were obtained and the rank-order coefficients of correlation of these ranks with the ranks by shop grades

3/ Appendix S

were calculated, Table 8. The rank-order coefficients of correlation for the B-200 group were 0.51, 0.39 and 0.43 for the three shortened forms, respectively, as compared with a rho of 0.41 for the whole test. The difference, 0.10, between the rho for the original test and the form shortened by biserial r was not significant. The CR of the difference was only 0.36 while the table required 2.12 for the .05 level of significance. No other shortened form of the test showed any superiority over the original test when compared with pupils' grades by the rank difference method of correlation.

Table 8.--COMPARISON OF CORRELATIONS OF PUPILS' RANKS BY GRADES AND ACCURACY AND FINISH SCORES ON WHOLE TEST AND VARIOUS FORMS ABBREVIATED BY REMOVING THE LEAST VALID ITEMS

School Groups	No. of Cases	Rank-Difference Correlations with Pupils' Grades			
		Whole Test	Shortened Forms of Test		
			r_{bis}	Clark	Lindquist and Cook
1	2	3	4	5	6
B-200	17	.41	.51	.39	.43
B-400	12	.12	.16	-.01	.06
D-200	20	.39	.18	.28	.30
D-400	13	.71	.60	.28	.71

When the weight of the time score was varied and the resulting combined scores for speed and quality correlated with pupils' grades, Table 9, the rho of the test scores of the B-200 group which were obtained from the use of "Mean time x 3" was 0.63. This was the highest rho for this group. For the B-400 group, the highest rho, 0.65, was obtained by using "Mean time" in the formula. The highest rho for the D-200 group, 0.32, was obtained with "Mean time x 2". For the D-400 group both "Mean time x 4" and "Mean time x 3" gave rho's of 0.70 which were high for that group.

Table 9.--RANK-DIFFERENCE CORRELATIONS OF COMBINED SCORES FOR QUALITY AND SPEED WITH PUPILS' GRADES, USING VARIOUS WEIGHTS FOR SPEED

School Groups	No. of Cases	[1] Mean time x 4	Mean time x 3	Mean time x 2	Mean time
1	2	3	4	5	6
B-200	17	.54	.63	.54	.52
B-400	12	.32	.43	.51	.65
D-200	20	.16	.17	.32	.02
D-400	13	.70	.70	.65	.56

[1]In Hunter's formula for the time score the denominator was "Mean time x 4," etc. The least weight for the time score was given by "Mean time x 4" and the greatest by using "Mean time" as the denominator.

Reliability

The reliability coefficients for the test, Table 10, found by various methods and for different groups of cases, varied from 0.795 to 0.903. For all 95 cases tested, the r's were from 0.824 by the odd-even split-half method and 0.894 by the method of rational equivalence. The difference between these two r's was 0.070. This difference had a standard error of 0.045 which gave a CR of 1.56. The difference was not significant. For the 31 cases that were retested, the smallest r was 0.795 which was found by the method of rational equivalence and the largest was 0.903 which was obtained by the odd-even method. The difference between these two r's was 0.108 which again is not significant. The CR is 1.44. For these same 31 cases, the reliability of a shortened form of the test, again through use of the odd-even split-half method, was 0.907.

Table 10.--RELIABILITY COEFFICIENTS OF THE TEST AS FOUND BY VARIOUS METHODS, BASED ON THE QUALITY SCORES

No. of Cases	Equivalent Halves: Experts' Division of Items	Equivalent Halves: [1]Item Difficulties	Odd - Even	Rational Equivalence	Retest
1	2	3	4	5	6
95	.838	.832	.824	.894	
31	.863	.883	.903	.795	.851
31	[2](Shortened form)		.907		

[1]The item difficulties are those shown by the performance of 95 pupils. The items were divided 1, 4, 8, 9, etc. for one half of the test and 2, 3, 6, 7, etc. for the other half of the test.

[2]The shortened form of the test was made by ignoring the performance on the eight items having the lowest biserial r's.

Chapter V

DISCUSSION

Chapter II, The Review of Literature, sought to present material from previous studies and investigations that might contribute to the answers to the questions in this study. Chapter IV, The Analysis of Data, presented a statistical analysis of the data that had been obtained by administration of the machine shop performance test to a total of 95 senior high school pupils. In the present chapter the bearing of this material on the questions in the problem analysis will be discussed. It should be emphasized that this is a study of the technique or method of development of a machine shop performance test and not the development of any specific test. Interest will be centered on the question "How?" rather than "What?"

What has been done in the field of performance testing

The answer to this question in Chapter II serves as a background and a starting point for this study. The actual number of performance tests available in print is very small, not many more than a dozen tests in all.

Machine shop performance tests.--The two performance tests in machine shop developed by Hunter (11,12) were for sophomores and juniors in a vocational school. The level of skill involved in these two tests is comparable to that developed in senior high schools that offer 400 periods of work in machine shop. The two machine shop performance tests published by Chapman (5) were designed by the army to classify new personnel as to their advancement in the trade. The performance test obtained from the Board of Examiners of the City of New York 1/ was employed to classify candidates for teaching licenses.

Other performance tests.--Chapman published eight performance tests besides those for machinists. They were for other mechanical trades and stenographers and typists. The Board of Examiners of the City of New York 2/ listed 22 trades -- mechanical and other -- for which licenses were granted. Holsinger (10) published a test in elementary electricity. Newkirk and Greene (16) published several performance tests for industrial arts classes. These included one in woodwork which had been administered to 30 pupils.

The place of performance tests.--Performance tests were considered desirable and useful. They were considered more reliable than grades or ratings ordinarily made by shop teachers. Authors who made generalizations such as Greene, Jorgensen and Gerberich (9) and Newkirk and Greene (16) stated that the performance test is feasible

1/ Appendix B

2/ Appendix B

but that up to the present time very little has been done.

<u>What operations should be included in a performance test for senior high school machine shop</u>

The answer to this question may be divided into three parts. The first deals with the criteria of a good performance test operation, the second with the selection of operations based on expert opinion and the third deals with the statistical validity of the individual items.

<u>Criteria of a good item for a machine shop performance test</u>.--The answer to this section should be sought in the review of literature.

Chapman (5) and Toops (22), in their discussions of the trade tests used in the army during the first world war, offered six criteria of a good item. These criteria have been reworded here to apply to high school machine shop classes. They are:

1. The operation must be representative of the work done in the class.
2. The operation must be one that can be performed by anyone in the class.
3. The operation must discriminate between pupils of different total achievement.
4. The operation must be easy to check or rate by mere measurement.
5. The operation must be easy to administer.

6. The operation must be one that can be performed with simple equipment.

Selection of operations.--Newkirk and Greene (16) Vocational Teacher Training Staff of the University of Pittsburgh (19) and Orleans (18) maintained that the content of a performance test -- that is, the operations -- is based primarily on expert opinion in the first selection. The operations in the test which is the vehicle for this study were selected cooperatively by nine senior high school machine shop teachers.

Item validity.--Item validity of operations in a machine shop performance test may depend on the nature of the item itself, which goes back to the selection of items covered in the above paragraph. It may also depend on the adequacy of the methods of inspection.

Since this study is one of technique, various methods of obtaining indices of validity for the items were investigated and compared. The biserial r was taken as a standard since it was recommended by Garrett (8) and Smith (21). The I.V's that resulted from the use of Clark's (6) method had a rank-difference correlation of 0.772 with the biserial r's. The correlation of the I.V.'s obtained with Lindquist and Cook's (14) method had a rank-difference correlation with the biserial r's of 0.880. A study of Lindquist and Cook's method shows that it involves only a fraction of the time needed to calculate the biserial r's for each item in the test. This fact plus the significant correlation of the two methods, 0.880. indicates that Lindquist and Cook's method may safely be used by industrial arts teachers in the investigation of the validity of operations in a performance test.

What controls must be exercised in administering the test

The controls are a definite part of the test as far as the teacher is concerned. According to Chapman (5) the controls should cover tools and equipment, the test administrator, the pupil and scoring.

The number, arrangement and condition of tools and equipment must be standardized. The teacher should know the directions for the pupils well enough to give them without hesitation and must not deviate in the least from the prepared directions.

The controls relating to scoring must cover the system of tolerances employed and also the specific directions for inspection of the test pieces. As for tolerances, the only actual comparison made indicated that the use of fixed tolerances as in industry, or of a sliding scale of credits as in some school shops, gave approximately the same results in ranking persons tested. In the study of inspection methods, the item receiving the lowest teachers' rating of 0, 1/16 chamfer, was also ranked low in inspection consistency and biserial r, being 30 and 29 respectively. The three items receiving teachers' ratings of 4, on a scale of 8, were 1/4, 30° chamfer and 1/4-20 thread square. They ranked 28, 3 and 26 by biserial r, and 13, 27.5 and 27.5 by inspection consistency. This coincidence of low rating of inspection methods by teachers, low rank in inspection consistency and low biserial r ranking suggests a relationship between methods of inspection adjudged unsatisfactory and low item validity. This further indicates the necessity of review of the methods of inspection after they have been tried out experimentally.

Toops (22) maintained that the use of gages is superior to measurement with scales and micrometer as a method of inspection.

What procedures must be used to determine the validity of the test

One common method of obtaining a statistical estimate of the validity of a test is to correlate it with a criterion. That method was employed in this study. The regular shop grades were employed as a criterion. The investigations of Newkirk and Greene (16) throw some doubt on the dependability of shop grades, but for the present there is no other criterion available. In the actual correlations between grades and quality scores the rank-order coefficient for the D-400 group was 0.71 which was significant. All the other coefficients were positive but too low to be significant. In the correlation of speed alone and grades, the coefficient for the B-400 group, 0.79, was significant.

In this category there were two negative coefficients. There were two significant coefficients among the correlations of shop grades and combined scores for quality and speed. They were 0.63 for the B-200 group and 0.70 for the D-400 group. The other three group coefficients were positive. These coefficients of correlation indicate two things: first, a performance test for speed alone would not be valid for the evaluation of pupil progress; second, a performance test for quality and speed combined is a more valid tool than either speed or quality alone.

Items of low item validity.--When the test was shortened by ignoring the inspection on certain items of low item validity no increase in the correlations with shop grades was noted. In only two cases were the coefficients increased and those increases were not significant.

The weight for time.--Throughout the study Hunter's (12) formula for time has been employed. Hunter arrived experimentally at a constant of 3 in his formula. The tests were rescored using constants of 4, 2, and 1. The resulting coefficients of correlation of combined scores for quality and speed with pupils' grades had the high coefficient for each school group either in the 3 constant or the 2 constant column. There was one exception, the coefficient of correlation for the D-400 group was the same for the constant of 3 as for the constant of 4. These higher correlations for the constants 2 and 3 support the soundness of Hunter's formula using 3 as the constant.

What does the test measure.--In all cases but one the differences between average scores of different school groups for quality alone and for quality and speed combined were significant. In the exception the difference was nearly significant. On the other hand, the differences between groups at different levels of instruction within the same school were much less significant. In only one case in four was the difference significant, and then barely so. These two facts seem to indicate that the pupil does not add appreciably to his manipulative skill during the second 200 periods of instruction.

What procedures must be used to determine the reliability of the test

Three methods of dividing the test into split-halves were employed. Two methods used what have been termed "equivalent halves". One pair of equivalent halves was based on expert opinion. The other pair was based on the difficulty of items as indicated by the results of pupils' performance. The third method was to divide the items on inspection by the odd-even method. Thirty-one cases were retested. Estimates of the reliability were also obtained by the method of rational equivalence. These methods gave four reliability coefficients for 95 cases, and five for the 31 retest cases. When these coefficients were compared it was found that there were no significant differences between them. This non-significance of the differences indicates that any one of the methods is about equally dependable as a means of obtaining the reliability coefficient of a performance test.

Suggestions for further study

No extensive investigation has been carried out on the question of the use of fixed tolerances or of sliding scales. It may be that for school purposes the sliding scale is superior to the system of fixed tolerances as used in the present study.

It seems that a performance test might have real diagnostic values in relation to the instruction given to a class as a whole. Could it be used to give a teacher information about weak points in the program of instruction that he is following?

The performance test and other criteria such as intelligence, mechanical aptitude, and post-school success in related fields offer possibilities for further studies.

Chapter VI

SUMMARY

There are a number of elements in industrial arts teaching that are measurable. Manipulative skill is one of these elements. To measure the effectiveness of teaching in this respect we must have reliable tools of evaluation. At the industrial arts level practically nothing has been done toward the development of evaluation tools in machine shop work. The problem of this study seems to be one of the first steps in the development of proper evaluation tools. The problem is, "How can a performance test in senior high school machine shop be developed?"

The development of a performance test involves: (1) selecting operations or determining the content of the test, (2) controlling the testing situation and the method of inspecting or rating the test product, (3) determining and improving the validity of the test, and (4) determining and improving the reliability of the test.

The selection of operations, in the experimental stage, must depend on the opinion of experts. In this study the operations were first selected from three authoritative texts. Then the list was reduced to include only the operations on which the pupils who were to be tested received instruction. The final step in selection was done cooperatively by eight senior high school machine shop teachers. They selected 21 operations which they thought could be combined in a machine shop test problem. Seventeen of these operations were used.

The testing situation includes the equipment and tools that are to be used, the condition of the machines and tools, and instructions for the pupil. All of these instructions must be adhered to rigidly. The instructions for rating the test product must be made out in detail for each dimension or feature of the product that is to be considered. After the test has been used experimentally, the relationship of item validity, consistency of inspection and the subjective evaluation of the inspection methods should be studied in order to detect items the low item validity of which may be due to unsatisfactory inspection methods. In cases where low rating and low inspection consistency coincide with low item validity, the methods of inspection should be improved if possible.

The statistical validity of a performance test may be determined by correlation with a criterion. Pupils' grades may be used although they are of doubtful reliability. Attempts in this study to improve the validity of the test by eliminating items of low item validity proved fruitless. The weight of the time score, which was originally used for the test, seemed to be as valid as 1½ times the original weight and more valid than any further increase of the weight or any decrease of the weight.

The reliability of a performance test may be determined by any of the ordinary methods: retest, split-half or equivalent forms, as well as by the method of rational equivalence. The reliability of the test was not improved by eliminating items of low item validity from the test.

A P P E N D I X

APPENDIX

TABLE OF CONTENTS

APPENDIX

TABLE OF CONTENTS.--Continued

APPENDIX A

MASTER DATA SHEETS

MASTER DATA SHEET NUMBER ONE

CASE NUMBER	PERIODS OF INSTRUCTION	SCHOOL	TOTAL SCORE	GROUP TOTALS: LATHE	GROUP TOTALS: MILL	GROUP TOTALS: BENCH-DRILL	TIME-TOTAL	TIME-MILL SET-UP	SCORE-INSPECTION + TIME	EQUIVALENT HALVES: A	B	C	D	SPLIT HALVES: ODD	EVEN	RETEST: TOTAL SCORE	TOTAL TIME	SCORE-INSPECTION + TIME	RANK
1	450	B	11				374		3.76	7	4	4	7	7	4	12	405		16/17
2	450	B	11				266		8.94	4	7	4	7	7	4	12	225		17/17
3	450	B	12				415		2.78	8	4	5	7	8	4	10	260		10/17
4	450	B	12				292		8.69	7	5	6	6	6	6	11	310		12/17
5	450	B	12				303		8.06	5	7	4	8	6	6	14	170		15/17
6	200	B	15				202		16.03	8	7	9	6	8	7				2/17
7	200	B	14				186		15.78	7	7	9	5	7	7				8/17
8	200	B	12				136		16.18	8	4	7	5	5	7	10	165		5/17
9	200	B	13				186		12.78	7	6	6	7	8	5				7/17
10	200	B	10				167		12.69	5	5	6	4	4	6	11	195		11/17
11	200	B	13				173		17.40	8	7	8	7	8	7				6/17
12	200	B	18				190		21.50	10	8	9	9	7	11				3/17
13	200	B	14				155		17.26	9	5	8	6	6	8	17	165		9/17
14	200	B	16				251		14.66	7	9	9	7	3	13				13/17
15	200	B	21				157		24.17	11	10	13	8	10	11				7/17
16	200	B	8				304		4.11	3	5	5	3	4	4				14/17
17	200	B	12				163		14.88	10	2	7	5	6	6				10/17
18	200	B	19				202		20.01	11	8	11	8	7	12				4/17
19	200	B	22				131		25.46	11	11	11	11	11	11	18	146		13/17
20	200	B	26				143		29.84	13	13	14	12	11	15	23	108		1/17
21	200	B	18				160		20.06	9	9	8	10	9	9				2/17
22	200	B	17				155		21.13	10	7	8	9	9	8				16/17
23	180	B	7				350		0.42	3	2	3	4	4	3				12/14
24	180	B	17				225		16.90	7	10	8	9	8	9	10	190		3/14
25	180	B	10				200		11.10	5	5	3	7	6	4	10	215		11/14
26	180	B	11				260		9.22	6	3	6	5	6	5	10	220		7/14
27	240	B	6				185		7.82	2	4	2	4	2	4				4/14
28	240	B	14				179		16.11	6	8	5	9	7	7				1/14
29	180	B	15				172		17.45	9	6	5	10	8	9				6/14
30	180	B	13				181		15.02	8	5	7	6	6	7	12	155		2/14
31	180	B	9				255		7.46	6	3	3	6	4	5	8	200		9/14
32	190	B	8				307		3.97	4	4	3	5	4	4				14/14
33	180	B	11				207		11.77	6	5	4	7	5	6				10/14
34	240	B	10				184		11.87	6	4	4	6	7	3				5/14
35	180	B	15				286		11.98	10	5	7	8	9	6				8/14
36	180	B	11				220		11.14	6	5	7	4	7	4				13/14
37	400	D	19				130		23.46	11	8	10	9	11	8				7/13
38	400	D	21				180		22.38	10	11	10	11	14	11				10/13
39	400	D	25				175		22.30	14	11	13	12	13	12				8/13
40	400	D	22				155		27.26	12	12	13	11	11	13				5/13
41	400	D	25				170		27.54	10	15	15	10	11	14				4/13
42	400	D	23				160		26.02	12	11	11	12	11	12				6/13
43	400	D	26				175		28.30	13	13	12	12	13	13				9/13
44	400	D	22				165		23.82	12	10	12	10	12	10				13/13
45	400	D	26				165		28.79	11	15	15	11	12	14				2/13
46	400	D	29				110		34.42	12	16	16	13	15	14				3/13
47	400	D	31				160		33.06	16	15	16	15	16	15				1/13
48	400	D	22				115		26.18	12	10	11	11	9	13				12/13
49	400	D	21				170		23.54	11	10	11	10	10	11				11/13
50	200	D	21				200		22.10	13	8	12	9	10	11	20	220		10/21
51	200	D	20				230		19.66	12	8	9	11	11	9	23	190		18/21
52	200	D	24				230		23.66	13	11	13	11	12	12				13/21
53	200	D	21				200		22.10	12	9	13	8	10	11	18	185		20/21
54	200	D	23				165		25.78	14	9	12	11	13	10	18	190		16/21
55	200	D	23				125		26.74	13	10	12	11	12	11	23	170		17/21
56	200	D	23				175		25.30	13	10	13	10	13	10	21	190		11/21
57	200	D	26				220		26.14	12	14	13	13	13	13	19	193		21/21
58	200	D	20				215		20.38	12	8	11	9	11	9	18	210		19/21
59	200	D	29				185		30.82	15	14	15	14	15	14	25	210		12/21
60	200	D	22				205		24.86	12	12	11	13	12	12				9/21
61	200	D	26				195		27.34	15	11	13	13	13	13				14/21
62	200	D	20				160		23.02	10	10	9	11	8	12				15/21
63	200	D	23				265		20.98	12	11	11	12	13	10	29	225		2/21
64	200	D	25				230		24.66	12	13	12	13	11	14	24	185		8/21
65	200	D	24				220		24.14	12	12	12	12	14	10	22	180		3/21
66	200	D	22				180		26.06	13	11	13	11	11	13	24	235		5/21
67	200	D	27				205		27.86	14	13	13	14	13	14	18	205		4/21
68	200	D	22				220		24.14	14	10	14	10	12	12	18	205		6/21
69	200	D	25				150		28.30	12	13	12	13	15	10	18	165		7/21
70	400	B	13				292		9.69	8	5	6	7	8	5				11/19
71	450	B	21				224		20.95	12	9	9	12	10	11				11/19
72	450	B	17				279		14.31	10	7	9	8	11	6				9/19
73	450	B	20				196		21.30	11	9	10	10	10	10				2/19
74	450	B	18				182		19.97	11	7	8	10	9	9				2/19
75	450	B	10				304		6.11	7	3	4	6	4	6				16/19
76	450	B	21				290		17.78	11	10	10	11	10	11				13/19
77	450	B	11				153		14.41	5	6	4	7	3	8				1/19
78	450	B	8				213		8.48	4	4	3	3	4	4				13/19
79	450	B	16				240		15.18	10	6	11	5	7	9				16/19
80	450	B	12				237		11.33	7	5	6	6	5	7				6/19
81	450	B	17				305		13.06	11	6	9	8	8	9				19/19
82	200	B	10				166		12.74	5	5	2	8	5	5				3/22
83	200	B	8				136		11.22	7	1	5	3	3	5				9/22
84	190	B	10				268		7.84	6	4	7	3	4	6				15/22
85	190	B	8				263		6.08	5	3	3	5	5	3				18/22
86	245	B	12				250		10.70	7	5	5	7	8	4				1/22
87	190	B	9				273		6.60	5	4	5	4	5	4				6/22
88	190	B	15				227		14.81	7	8	7	8	7	8				4/22
89	200	B	6				190		7.58	5	1	3	3	4	2				13/22
90	150	E	20				454		9.91	11	9	10	10	11	9				3/14
91	150	E	17				449		8.15	9	8	10	7	8	9				12/14
92	150	E	13				413		3.88	8	5	4	9	7	6				10/14
93	150	E	17				216		7.74	8	9	9	8	8	9				5/14
94	150	E	21				371		13.90	9	12	12	9	10	11				4/14
95	150	E	19				410		10.02	8	11	11	8	9	10				11/14
96																			
97																			
98																			
99																			
100																			

INSPECTION ITEMS (FINISH) — ITEM TOTALS:

ITEM TOTALS	44	55	70	49	61	90	61	35	48	56	35	30	23	49	40	81	42	66	80	38	71	44	28	46	42	31	44	49	34	46	31	38

MASTER DATA SHEET NUMBER TWO

CASE NUMBER	2	4	11	12	13	14	15	16	20	21	22	23	24	25	26	28	29	30	37	46	50	51	58	64	90	91	92	93	94	95
SCHOOL	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	O	O	O	O	O	O	E	E	E	E	E	E

INSPECTION ITEMS — ITEM TOTALS

	T	R1	R2
4 1/4	16	21	21
3	17	21	20
1 1/4 THREAD	26	26	26
3/4	19 / 19	19	18
1/4	18	21	20
31/32	28	30	30
11/16	18	22	19
1/16 CHAMFER	14	12	11
3/32	13	14	14
.188	14	12	10
.750	17	22	21
THREAD P.D.	11	11	11
.075	9	8	8
1 1/4 HOLE	17	15	15
3/8 REAM	10	23	25
12°	24	27	27
30°	10	11	10
90° HOLES	19	23	23
¼-20 DEPTH	25	23	25
¼-20 SQUARE	12	11	9
1/64 RADIUS	22	22	23
SHOULDER-SQUARE	15	18	22
CENTERING HOLE	7	5	5
CENTERING KEYWAY	11	16	15
END FLAT	10	10	10
FINISH 12° ANGLE	11	8	9
FINISH 31/32 DIA	13	12	14
FINISH .750 DIA	16	15	19
FINISH THR. END	13	15	16
FINISH SHOULDER	17	16	15
FINISH TAP. END	21	18	18
BURRING	11	5	8
CASE TOTALS			
RANKS			

Machine Shop Performance Test
Criticism of Instructions for the Inspection of Test Pieces

Dimension	Teachers' rating	[1]Composite rating
4½ length	- + + +	6
3 length	0 + + 0	6
1¼ length of thread	+ + + +	8
3/4 to 5/16 drilled hole	+ - + 0	5
¼ to ¼-20 tapped hole	- - + +	4
31/32 diameter	+ + + +	8
11/16 diameter	+ 0 + +	7
1/16 chamfer	- - - -	0
3/32 depth of keyway	+ 0 + 0	6
.188 width of keyway	+ + 0 +	7
.750 diameter	0 + + +	7
Thread pitch diameter	+ 0 + +	7
.075 depth of undercut	0 0 0 +	5
1¼ depth of reamed hole	+ + 0 +	7
3/8 diameter of reamed hole	0 + + +	7
12° angle, turned	0 0 + +	6
30° angle of chamfer	+ - - +	4
90° relationship of the 5/16 drilled hole and the ¼-20 tapped hole	0 + + +	7
¼-20 thread, depth	0 + + +	7
¼-20 thread, squareness	- + + -	4
1/64 maximum radius	0 + + +	7
Shoulder, squareness	+ + + +	8

[1] 0 for - , 1 for 0, 2 for + .

MASTER DATA SHEET NUMBER THREE

Machine Shop Performance Test
Criticism of Instructions for the Inspection of Test Pieces--continued

Dimensions	Teachers' rating	Composite rating
Centering, 5/16 drilled hole	- + + 0	5
Centering, keyway	+ + 0 0	6
End, flatness of thread	+ 0 + +	7
Burring all edges	+ + + +	8
Finish of turned surfaces	+ + + +	8
Finish of faced surfaces	+ + + +	8

APPENDIX B

LETTER FROM DR. E. J. GANNON, AND ENCLOSURE

Board of Education
of the City of New York
THE BOARD OF EXAMINERS
110 Livingston Street, Brooklyn 2

4 February 1949

Mr. Richard B. Smith
193 Morningside Drive East
Bristol, Connecticut

My dear Mr. Smith

I enclose a list of trade subjects, instructions to assistant examiners, a rating sheet, and a sample job sheet for a performance test in machine shop work.

I trust that this material will be of assistance to you.

Very truly yours

Edmund J. Gannon
Chairman, Committee on Industrial Licenses

EJG
JED/D

(Enclosure)

Shop Subjects Trades

Air conditioning (M)(T)
Baking (M)
Beauty culture (W)
Building maintenance (M)
Cafeteria and catering (M)
Cafeteria and tea room training (W)
Clock and watch mechanics (M)
Commercial and domestic refrigeration (M)(T)
Commercial photography (M)
Dental mechanics (M)(T)
Electrical installation & practice (M)
Floristry (M)(W)
Gas and electrical welding (M)
Jewelry making (M&W)
Leather goods manufacturing (M)
Maritime trades (deck) (M)
Maritime trades (engine) (M)
Maritime trades (steward) (M)
Mechanical drafting (M)(T)
Radio mechanics (M)
Trade dressmaking (W)
Women's & children's garment manufacturing (M&W)

Board of Examiners | Board of Education

Performance Test
Job Sheet

Examination for License to Teach Machine Shop Work in Day High Schools

September 27, 1941 | Time: 6 hours--8:30 A.M. to 2:30 P.M.

Name______________________ Bench No.______ Candidate's No.____

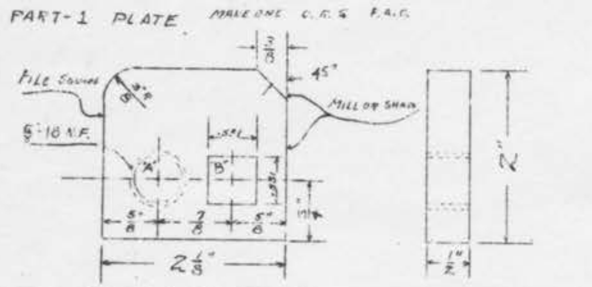
Address______________________ Work began__________
Work stopped__________

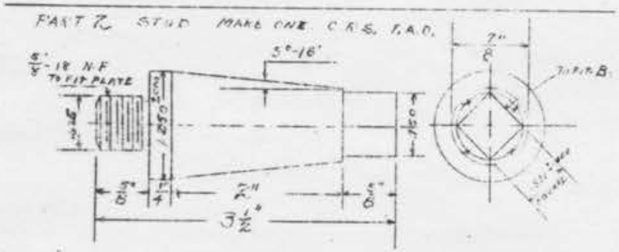
Note: This job sheet is to be kept on your bench until the close of the session.

You are to do the following:

1. Sign the job sheet, the rating form, and the accident release. The job sheet is to remain on your work bench. The other papers will be collected promptly.
2. Examine the stock given to you. If it is not satisfactory for the job, it may be exchanged now. After you begin work, no exchanges will be made nor will second trials be permitted.
3. At the close of the test, hand in to the examiner your work and the job sheet.
4. Candidates may be excused from the session after the first hour, with the approval of the examiner-in-charge.
5. All diameters must be concentric with the stud.
6. Conversation among candidates is not permitted.

PROBLEM: MAKE THE PLATE AND STUD SHOWN IN THE SKETCH BELOW.





Board of Examiners | Board of Education

Rating Sheet

Performance Test

Examination for License to Teach in Day High Schools

Machine Shop Work

September 27, 1941 | Time: 6 hours--8:30 A.M. - 2:30PM

Bench No.________ Candidate's No.______

Candidates are advised that no appeal will be entertained from the findings of this performance test. If the conditions attending the test are unsatisfactory, candidates are to call the attention of the examiner at the time of the test to such conditions. Candidates who are not in sufficiently good health to take a test at the time they are summoned are directed to inform the examiner of this fact; otherwise they are to sign the following:

I affirm that I am in sufficiently good health to take this examination and that I am not indisposed physically or mentally.

(Signature) ______________________

PROBLEM: MAKE THE PLATE AND STUD SHOWN ON THE JOB SHEET.

Elements to be rated	Unsat. 0	1	2	Sat. 3	4	5	Comments
PART I							
1. Accuracy of layout.							
2. File and mill ends square.							
3. Drill and tap hole.							
4. File square hole.							
5. Bevel and round corners.							
PART II							
6. Accuracy of lengths of stud.							
7. Thread cutting to fit plate.							
8. Shoulder turning and outside diameters.							
9. Taper turning.							
10. Mill square shank to fit plate.							

Maximum: 50 credits ________%

Was job completed? Yes No (Cross out one)

REMARKS AND COMPENSATORY SKILLS, etc.

Place of Test:______________________ Examiners: ______________

D. ______________

APPENDIX C

LETTER FROM R. S. HUNTER, IN PART

(copy)

145 E. Idaho St.
Boise, Idaho.
Sept. 15, 1947

Mr. Richard B. Smith
193 Morningside Drive East
Bristol, Conn.

Dear Mr. Smith:

Your letter of Aug. 25th sent to Pittsburgh was forwarded to me -- hence the delay in answering it.

I was interested in your inquiry concerning the discrepancy in the formula published in 1942 and the one used for the study in 1945. You are quite correct in your observation, namely, that the original formula placed too much emphasis upon speed -- at least for beginning students. In particular, it imposed a severe handicap upon the slow but accurate students. It may be conceded that the slow but accurate student should be penalized, but he should not be penalized to the extent that he tends to become discouraged, or for that matter, that he gains in speed at the expense of accuracy. Accordingly, I arbitrarily used the factor of 3 to reduce the weighting of speed. Later I checked the factor in order to determine whether the weighting tended to decrease accuracy or increase it. I was interested in maintaining accuracy.

....

The revised formula appears to perform satisfactorily for student training purposes. Furthermore, it correlates rather well with another investigator's grading system.

I trust that I have answered your question to your satisfaction.

Cordially yours,

Robert S. Hunter.

APPENDIX D

MACHINE SHOP PERFORMANCE TEST, MANUAL FOR EXPERIMENTAL USE

MACHINE SHOP PERFORMANCE TEST

Manual for
Experimental Use
1949

Prepared by

Richard B. Smith, Head of Machine Shops, Bristol (Conn.) High School

In cooperation with

John Armen, Norwich Free Academy, Norwich, Conn.

Henry W. Bartnikowski, Stamford (Conn.) High School

Byron A. Berry, Bristol (Conn.) High School

Edward D. Busby, Bristol (Conn.) High School

George M. Crook, Pine Manual Training School, Ansonia, Conn.

T. O. Kennedy, Norwich Free Academy, Norwich, Conn.

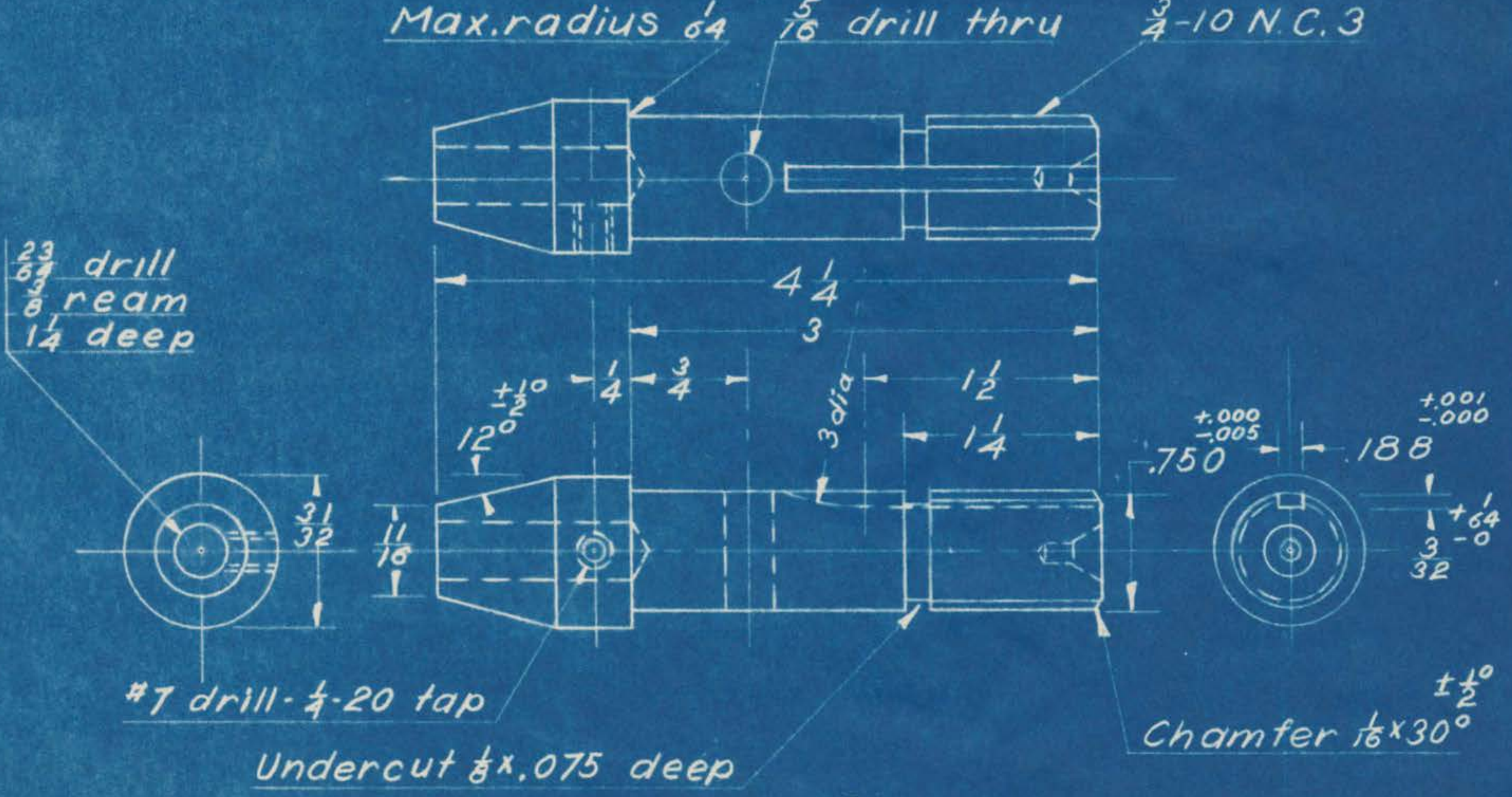
Joseph Madrick, William Hall High School, West Hartford, Conn.

John F. Shea, Weaver High School, Hartford, Conn.

Contents

Machine Shop Performance Test Problem	Material CRS 1 dia x 4 5/16 lg.	Date: 5/2/49 Drawn by: RBS	Scale 1=1

Other fract. tolerances ± 1/64
Other deci. tolerances ± .005



Burr all corners

Case No. ___

Machine Shop Performance Test
Individual Record

Name of pupil ______________________ School ______________

Class in school ____ No. periods (45 min.) instruction ______ Number in class ___

Rank in class _____ Date of test __/__/__ to __/__/__

Name of test administrator ______________________

. . . .

INSPECTION RECORD

Dimension	Tolerance	Meas.	OK
$4\frac{1}{4}$	± 1/64		
3	± 1/64		
$1\frac{1}{4}$ - thread	± 1/64		
3/4	± 1/64		
1/4	± 1/64		
31/32	± 1/64		
11/16	± 1/64		
1/16 - chamfer	± 1/64		
3/32	+ 1/64 − 0		
.188	Gage	-	
.750	+ .000 − .005		
Thread P.D.	+ .000 − .004		
.075	± .005		
$1\frac{1}{4}$ - hole	Gage	-	
3/8 - ream	Gage		
12° angle	± 1/2°		
30° chamfer	± 1/2°		
90° - holes	Visual	-	
$\frac{1}{4}$-20 thread	Depth	-	
$\frac{1}{4}$-20 thread	Square	-	
1/64 radius	Max.	-	
Shoulder	Square	-	
Centering-hole	Gage	-	
Centering-keyway	Gage		
End	Flat	-	
Burring all edges		-	
Finish on turned-faced surfaces			
12° angle			
31/32 diameter			
.750 diameter			
Facing-threaded end			
Facing-shoulder			
Facing-tapered end			

RECORD OF TIME WORKED ON THE TEST

Date	Start Time	Breaks			Stop Time	Minutes Worked
		Start	End	No. Minute		

Total minutes worked ____

Time on milling machine

Start-set-up time ____:____

Finish-set-up time ____:____

Milling set-up time-in minutes ____

Total work time less milling machine set-up in minutes ____

Total number of items correct on inspection ____

Score, corrected for speed (Hunter's Formula) ____

Machine Shop Performance Test
Instructions to Test Administrators

Preparations for the test

1. Familiarize yourself thoroughly with the various parts of the test:

   a. The blueprint, so that you can answer readily any question about it

   b. The "Instructions to Pupils", so that you can read them freely

   c. The "Instructions to Test Administrators" for "Administering the Test"

   d. The "Individual Record"

2. Tools:

   See that the tools required for the test are in their places in the tool crib or the tool panels. Check and sharpen, if necessary, drills, reamers, taps, thread-cutting and undercutting tools. See that crib attendants, if any, are ready to give prompt service to those taking the test.

3. Machines:

   a. Lathes:

      Check the centers for running truth in the headstocks and check and adjust the tailstocks for taper as specified.

   b. Milling machine:

      Have the speed and feed set as specified. The vise should be set up and indicated. The spindle of the milling machine should be clear.

   c. Drill press:

      Prepare as specified.

4. Cut enough pieces of stock for one piece for each pupil. The stock should be as long as specified and not more than 1/16" longer.

5. Prepare a tag for each pupil's work with his name clearly written on it.

6. Prepare an "Individual Record" for each pupil. Write his name on it.

7. Read the "Instructions to pupils about preparation for the test" to the pupils. This should be done at least one day before the test so that they can take care of their tool bits. The question of steel rules is to be left to the normal procedure in each shop.

Machine Shop Performance Test
Equipment for the Test

Machines and accessories

1. Lathes:

9" Model A South Bend screw cutting lathe or its equivalent or better in good running order. The dog plates should be mounted on the spindles. The lathes should be equipped with two good centers which can be removed and replaced without altering the taper. The live center should run true within .001". (.002" by indicator) The lathes should be corrected for taper to .001" or better in 4" of length using a $4\frac{1}{4}$" test bar. The compound rest should be set at 0°. The lathes may be equipped with thread chasing dials.

Accessories: (For each lathe, arranged within easy access of the machine)

3 jaw universal chuck, accurate to within .006" (.012 by indicator)
Key for the above chuck
Drill chuck for the tailstock, 3/8" capacity or larger and chuck key
Straight and left hand tool holders
Lathe dogs
Wrenches for the tool posts, compound rests and tailstocks

2. Milling machine:

Size adequate to take the full cut on a 3/16" x 3/32" keyseat at a feed of $3\frac{1}{4}$" per minute. The spindle should be set at a speed of approximately 50 R.P.M. and the feed at approximately 2" per minute.

Accessories: (Arranged within easy access of the machine)

3/16" saw or side-cutting plain cutter, 3" diameter preferred
Arbor for the above cutter and a draw-bar
Wrench for the draw-bar
Plain vise, 4" jaw or larger
Vise handle
Lead hammer
Brass faces for the vise jaws
Parallel, approximately $\frac{1}{2}$" less in height than the vise jaw.

3. Drill press:

3/8" capacity or larger with a spindle speed range of at least 700 - 1400 R.P.M.

Accessories: (Arranged near the drill press)

Chuck key
Drill press vise, plain jaws, 3" or larger
Vise handle, if needed

4. General

Cutting oil
White lead or other lubricant

Hand and small tools

1. Measurement and layout:

    Center head with blade
    Scriber
    Hermaphrodite calipers
    Layout fluid or chalk
    Punches
    Hammer
    Combination square, 12"
    Combination square, 6"
    Center gage
    1" micrometer
    4" outside calipers
    Thread ring gage, 3/4 - 10, NC-3
    Screw pitch gage
    Protractor
    Feeler gage

2. Lathe:

    Combination drill and countersink
    23/64 drill
    3/8 machine reamer
    10" mill file, second cut or smooth
    Thread-cutting tools
    Undercutting tools, 1/8"
    Brass strip for chuck jaws

3. Drill press and bench:

    #7 drill
    5/16 drill, sharpened to drill not more than .004" oversize
    1/4 - 20 tap
    Tap wrench
    V-block and clamp
    10" mill file, second cut or smooth
    Triangular scraper

4. Each pupil's personal equipment:

    Lathe tools for turning and facing (It is preferred that these lathe tools be those for which the pupil is normally responsible)
    Steel rule, preferably 6" (The pupils may use those from the tool crib if that is the custom of the shop)
    Pencil

Machine Shop Performance Test
Instructions to Test Administrators

Administering the test

1. Start only as many pupils at one time as your equipment will allow to work without an undesirable amount of waiting for machines, tools, etc. You will need one milling machine and one drill press for each three lathes.

2. Watch carefully to eliminate any help from other pupils.

3. Start each pupil on the test by reading to him the "Instructions to pupils at the start of the test. These instructions must be followed absolutely so that the results of the test will be consistent from school to school.

4. The use of the "Individual Record" - Record of time worked on the test

   a. Record the starting time, break times and stopping time for each pupil each day. "Break time" is time lost by a pupil due to inability to obtain a proper machine or tool. If there is more than one break in any day record the second break directly under the first one.

   b. At the end of each day draw a horizontal line under the time record for that day.

   c. Record the time that the pupil goes to the milling machine as "Start-set-up time". Record the time that the pupil completes the set-up of the cutter and arbor as "Finish-set-up time". The purpose of this time record is to see if the inclusion of the milling set-up in the test is desirable.

5. At the end of each day's work collect the test pieces and tag them immediately. Collect the blueprints.

6. The lathes should be checked from day to day to see that they continue to meet the specifications.

7. The milling machine must be returned to its original specifications after each pupil has used it. For the last two items see "Equipment for the Test", page 5.

Machine Shop Performance Test
Instructions to Pupils

Note to test administrators:

These instructions are to be read as given below and are not to be enlarged upon nor explained, but only repeated.

Instructions to pupils about preparation for the test, which should be given long enough in advance of the actual date of the test for the pupils to make the preparations suggested: (Omit the reference to steel rules if it is the practice in your shop to issue steel rules from the tool crib)

"On (give date of test) you are to start on a performance test problem. In preparation for the test you are advised to see that your lathe tools for facing and turning are in good condition. You will be provided with thread cutting tools. You should also provide yourselves with 6" steel rules and pencils."

Instructions to pupils at the start of the test; hand the pupils the blueprints and say:

"You are to make this piece as a performance test. Examine your blueprints. (Pause a few seconds) You may ask questions about the blueprint later if you wish to. You will not be given any advice about the procedures to follow, and, since this is a test, you are requested not seek any help from other pupils.

" You will do your lathe work on the following lathes: (Indicate a lathe for each pupil) These lathes are free from taper to a degree satisfactory for this test and have good centers. You will use (Indicate which) milling machine for the milling operations. You will obtain your tools in the regular way. Undercutting and thread-cutting tools will be provided.

"You may check and oil your lathe before starting the test. Your time will start when you are handed your piece of stock. If you have to stop work at any time because of inability to obtain a needed tool or machine, inform me immediately so that the time you lose will not be counted as working time. Turn in your test piece as soon as you stop work at the end of the class and then clean up your machine and turn in your tools.

"The completed test will be scored for accuracy, finish, speed and correct interpretation of the blueprint. Speed usually counts about nine percent of the score. Finish counts about 1/5 of the score. Turned surfaces are not to be filed or polished. The 3/4 - 10 thread is to be finished on the lathe. If you should spoil the test piece as far as one dimension or operation is concerned, proceed with the test and complete the rest of it according to the blueprint. You will not be given a second piece of stock.

"Questions about the blueprint must be asked of me individually later. Do you have any other questions?"

In case of questions, reread the part of the instructions that applies to the question asked.

Instructions to pupils when they restart the test on a subsequent day:

"Set up your work in your machine just as it was when you stopped work on the test. Have me check it before you start work. If you are moving to another machine or are making a new set-up, your time will start when you receive your test piece. Do not start work until I tell you to do so."

Machine Shop Performance Test
Instructions to Test Administrators

Scoring and the "Individual Record"

1. Leave the Case No. blank.

2. For the class in school use a number. (9, 10, 11, 12)

3. For the number of periods of instruction, enter the total number of periods of instruction in machine shop that the pupil has had in any of the four years of high school. If the length of the period is more than five minutes longer or shorter than 45 minutes, cross out the "45" and write in the nearest multiple of five.

4. For the number in the pupil's class combine all sections of the same class.

5. The pupil's rank in his class should be based on his regular shop grades for at least the last four months. Pupils should be ranked as first, second, third, etc. in the class. Do not give the average grade.

6. The keeping of the time record is covered in the section on "Administering the Test".

7. The inspection record:

    a. Measure the first eight dimensions in the inspection record with a steel rule and calipers. Record the actual measurement to the nearest 64th in the measurement column. If the measurement is within the tolerance of plus or minus one 64th, check (✓) the OK column.

    b. Measure the 3/32" depth of the keyway by placing the .115" square stock in the keyway and measuring to the opposite side of the .750" diameter. Make the measurement 1/8" above the undercut. Record the actual measurement over the square stock. The correct measurement is from .759" to .774". If the measurement is within these limits, check the OK column.

    c. Measure the width of the keyway, .188", with the 3/16" square stock. The key should enter the keyway, but should show no play. If satisfactory, check the OK column.

    d. Measure the .750"diameter with a micrometer. Record the actual measurement. The correct measurement is from .745" to .750". If the measurement is within these limits, check the OK column.

    e. Measure the pitch diameter of the screw thread (3/4 - 10) with the .0625" wires and a micrometer. Record the actual measurement over the wires. The correct measurement is from .782" to .786". If the measurement is within these limits, check the OK column.

    f. Measure the .075" depth of the undercut by placing two pieces of the .115" square stock in opposite sides of the undercut. Measure over the square stock with a micrometer. Record the actual measurement. The correct measurement is from .825" to .835". If the measurement is within these limits, check the OK column.

g. Check the $1\frac{1}{4}$" depth of the reamed hole with the plug gage. The gage should go into the hole at least as far as the shoulder. If it does, check the OK column. Remove the burr where the tapped hole enters the reamed hole if necessary to make this check.

h. Check the 3/8" diameter of the reamed hole with the plug gage. The gage should go into the hole the full required depth without any play. If it goes in satisfactorily, check the OK column.

i. Measure the 12° angle and the 30° chamfer with a protractor. Record the measurements to the nearest $\frac{1}{2}^\circ$. If they are within the tolerance of plus or minus $\frac{1}{2}^\circ$, check the appropriate space in the OK column.

j. Check the 90° relationship of the $\frac{1}{4}$-20 tapped hole and the 5/16" drilled hole. If the tapped hole is up when the test piece is held with the reamed hole towards you and the keyway towards the left, check the OK column.

k. Check the depth of the $\frac{1}{4}$-20 tapped thread with a commercial screw or a plug tap. If there is a full thread into the reamed hole, check the OK column.

l. Check the squareness of the tapped hole. Screw the plug tap or commercial screw into the tapped hole. Check the squareness with a solid square. If no out-of-squareness can be detected, check the OK column.

m. Check the 1/64" max. radius with a 1/64" radius gage. If the radius is 1/64" or less, check the OK column.

n. Check the squareness of the shoulder with a 1/32" radius gage or a good steel rule. If the shoulder is perfectly square, showing no light under the rule or radius gage, check the OK column.

o. Check the alignment of the 5/16" drilled hole with the centerline of the work with the centering gage. If the gage goes over both sides of the .750 diameter with the pin in the hole, check the OK column. (This gage allows the hole to be .006" off the centerline.)

p. Check the centering of the keyway with the keyway centering gage. Place the gage in the keyway with the hook over the threaded end of the test piece. The end of the hook should be exactly in the center of the center hole. Hold the test piece and the gage with the hook of the gage vertical. Measure the distance from the end of the hook to each side of the center hole. Record the actual difference between these two measurements in 64ths. If the difference is not more than 1/64", check the OK column.

q. Check the burring of all sharp corners - 11/16", shoulder, 5/16" hole, 3/8" reamed hole. If all have been burred, check the OK column.

r. Check the finish of each turned and faced surface. Compare each turned surface with the turned sample. Compare each faced surface with the faced sample. If the surface being checked is as good or better than the sample, check the OK column.

g. Check the $1\frac{1}{4}$" depth of the reamed hole with the plug gage. The gage should go into the hole at least as far as the shoulder. If it does, check the OK column. Remove the burr where the tapped hole enters the reamed hole if necessary to make this check.

h. Check the 3/8" diameter of the reamed hole with the plug gage. The gage should go into the hole the full required depth without any play. If it goes in satisfactorily, check the OK column.

i. Measure the 12° angle and the 30° chamfer with a protractor. Record the measurements to the nearest $\frac{1}{2}^{\circ}$. If they are within the tolerance of plus or minus $\frac{1}{2}^{\circ}$, check the appropriate space in the OK column.

j. Check the 90° relationship of the $\frac{1}{4}$-20 tapped hole and the 5/16" drilled hole. If the tapped hole is up when the test piece is held with the reamed hole towards you and the keyway towards the left, check the OK column.

k. Check the depth of the $\frac{1}{4}$-20 tapped thread with a commercial screw or a plug tap. If there is a full thread into the reamed hole, check the OK column.

l. Check the squareness of the tapped hole. Screw the plug tap or commercial screw into the tapped hole. Check the squareness with a solid square. If no out-of-squareness can be detected, check the OK column.

m. Check the 1/64" max. radius with a 1/64" radius gage. If the radius is 1/64" or less, check the OK column.

n. Check the squareness of the shoulder with a 1/32" radius gage or a good steel rule. If the shoulder is perfectly square, showing no light under the rule or radius gage, check the OK column.

o. Check the alignment of the 5/16" drilled hole with the centerline of the work with the centering gage. If the gage goes over both sides of the .750 diameter with the pin in the hole, check the OK column. (This gage allows the hole to be .006" off the centerline.)

p. Check the centering of the keyway with the keyway centering gage. Place the gage in the keyway with the hook over the threaded end of the test piece. The end of the hook should be exactly in the center of the center hole. Hold the test piece and the gage with the hook of the

p_1. Check the flatness of the threaded end of the test piece with a steel rule or a machinist's square. No light should be visible except in the center hole. If satisfactory, check the OK column.

q. Check the burring of all sharp corners - 11/16", shoulder, 5/16" hole, 3/8" reamed hole. If all have been burred, check the OK column.

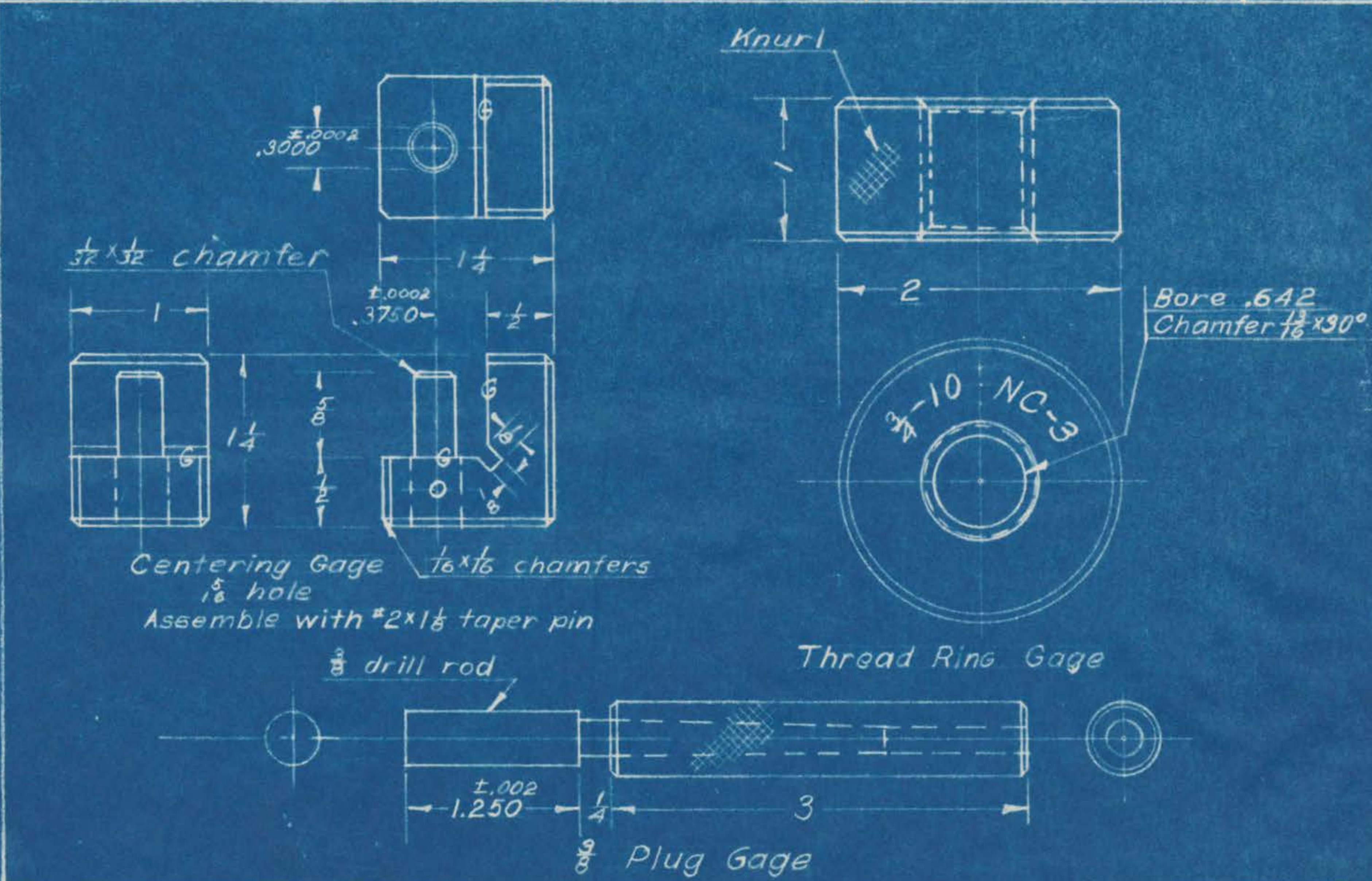
r. Check the finish of each turned and faced surface. Compare each turned surface with the turned sample. Compare each faced surface with the faced sample. If the surface being checked is as good or better than the sample, check the OK column.

8. Measurement and inspection tools:

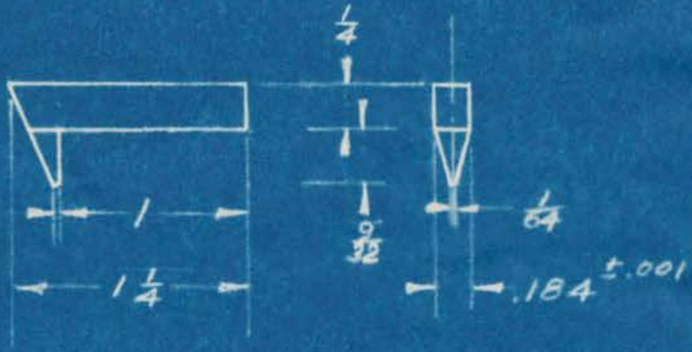
6" steel rule
4" outside calipers
1" micrometer
2 pc. .115" square stock*
1 pc. 3/16" square stock*
3 pc. .0625 wires*
3/8" plug gage*
Vernier protractor
1/64" radius gage
$\frac{1}{4}$ - 20 thread plug gage, or plug tap, or 2" long machine screws
Solid machinist's square
Centering gage for 5/16" hole*
Centering gage for keyway*
Turned surface sample*
Faced surface sample*

* To be furnished to the test administrator along with the blueprints of the test problem and copies of the "Individual Record".

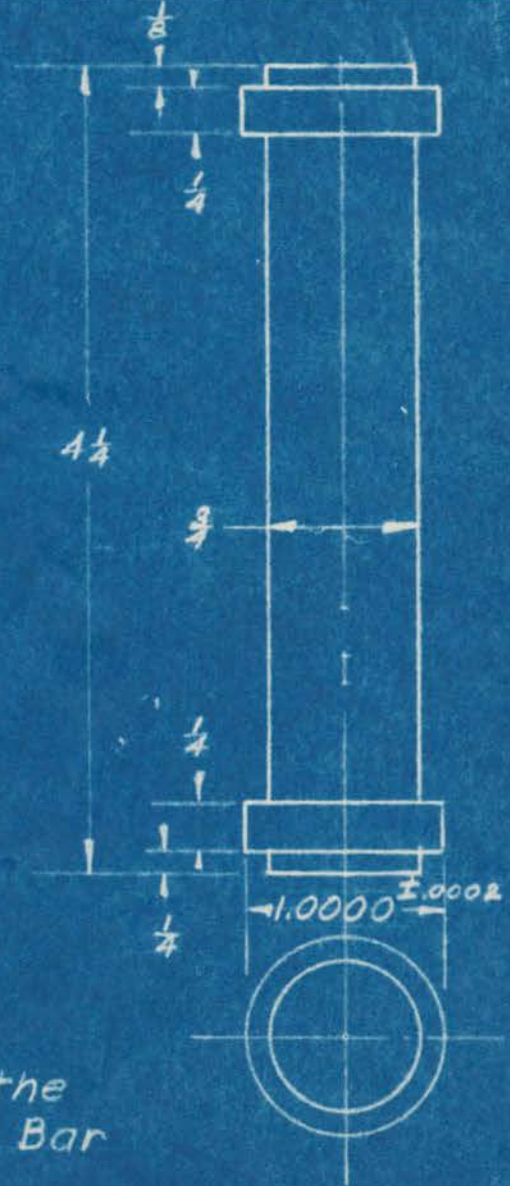
Machine Shop Performance Test Gages	Material CRS Drill Rod	Date: 14-2-49 Drawn by: RBS	Scale 1=1



Machine Shop Performance Test Gages	Material CRS	Date: 17-2-49 Drawn by: RBS	Scale 1=1



Centering Gage –
3/16 keyway
Note: The 1/64 point must be centered within .002 in relation to .184 shank



Lathe
Test Bar

APPENDIX E

LETTER TO CONNECTICUT

INDUSTRIAL ARTS TEACHERS

OF MACHINE SHOP

193 Morningside Drive East
Bristol, Conn.
September 22, 1947

Mr. G. Waldo Healy
Hall High School
West Hartford, Conn.

Dear Mr. Healy:

As a part of my work for a Master of Education degree at the Colorado A and M College, Fort Collins, Colo. I have to develop a performance test that can be used with high school pupils who have had from 200 to 400 periods (150 to 300 hours) of instruction in machine shop. A performance test is defined as one in which the pupil uses the tools and machines of the machine shop to alter the size, shape and surface quality of metal. The development of this test is to be the subject of my thesis.

In order to obtain an adequate number of pupils to be tested, I shall need the cooperation of some other machine shop instructors in determining certain elements of the test and in administering the completed test to a limited number of pupils towards the end of the school year.

I am writing now to ask if you would be interested in cooperating in the study at any stage of its progress. As I see the study now the various steps would be approximately as follows:

1. Check a list of machine shop operations, which I shall prepare, to determine those to be included in the test.

2. Check the instructions for administering the test to see if they can be followed in your shop.

3. Check the blueprint of the test problem.

4. Perform the test yourself as an expert to provide an analysis of the test elements.

5. Administer the test to a selected number of your pupils. Provide me with data relative to the grades of your pupils. (You may, of course, use the test for as many of your pupils as you wish for your own records.)

I am inclosing a typed form for you to indicate in which, if any, parts of the study you wish to participate.

Cordially yours,

Richard B. Smith
Machine Shop Department
Bristol High School

Copy to Mr. Harriman.

Mr. Richard B. Smith,
193 Morningside Drive East,
Bristol, Conn.

Dear Mr. Smith:

I shall be able to take part in the development of the machine shop performance test that you propose in the following ways:

1. Check the list of machine shop operations. ____
2. Check the instructions for administering the performance test. ____
3. Check the blueprint of the test problem. ____
4. Perform the test myself. ____
5. Administer the test to some of my pupils and supply data relative to their grades in machine shop. ____

I regret that I shall not be able to take part in the study at all.

_______(Check here)

Data:(Please supply now if possible)

1. Approximate number of machine shop pupils in 1947-8 ___
2. Approximate number of pupils who will complete 200 periods of machine shop instruction by the end of this year. ___
3. Approximate date these pupils will complete 200 periods of instruction. ___
4. Approximate number of pupils who will complete 400 periods of instruction in the machine shop by the end of this year. ____
5. Approximate date these pupils will complete 400 periods of machine shop instruction.

Signed___________________

Address__________________

____

Telephone, home__________school___________

APPENDIX F

CHECK FORM OF THE AREAS OF INSTRUCTION, AND SUMMARY OF REPLIES

CHECK OF AREAS OF INSTRUCTION AND SUMMARY OF REPLIES

Check List

I give instruction on the following machines in my machine shop: (Please write the number of each type in the space)

Engine lathes___ Milling machines___ Shapers___

Planers___ Drill presses___ Tool and cutter grinders___

Surface grinders___ Cylindrical grinders___

Others____

Signed________________________________

________________________High School

Summary of Replies

Name of machine	Schools					
	A	B	D	E	F	G
Drill press	3	2	2	[1]Y	2	3
Bench work		Y				
Engine lathe	10	20	6	Y	9	12
Milling machine	2	8	1	Y	1	2
Shaper	1	2	2	Y	2	1
Planer		1				
Surface grinder	1	1	1	Y	1	1
Cylindrical grinder						
Tool and cutter grinder		1				
Turret lathe		1		Y		

[1]Indicates a positive answer without giving the number of machines.

APPENDIX G

CHECK LIST OF MACHINE SHOP OPERATIONS

Checking Columns
A B

Bench operations

1. Lay out centers in round stock with a center head and scriber.
2. Lay out centers in round stock with hermaphrodite calipers.
3. Lay out centers in round stock with V-block, surface plate and surface gage.
4. Lay out centers in rectangular stock.
5. Lay out holes for drilling.
6. Lay out curves and tangents.
7. Scribe circles.
8. Lay out rectangular stock for milling.
9. Lay out lines with a surface gage.
10. Lay out lines with a vernier height gage.
11. Cut square or round stock with a hand hack saw.
12. Cut thin stock with a hand hack saw.
13. Grind a flat cold chisel.
14. Shear thin stock with a cold chisel.
15. Chip a flat surface.
16. Cut an oil groove with a chisel.
17. Chip with a power chisel.
18. B roach with a drift.
19. Cross file a flat surface.
20. Draw file a flat surface.
21. Sharpen a flat scraper.
22. Scrape a flat surface.
23. Scrape a bearing.
24. Drill with a breast drill.
25. Ream with a hand reamer.
26. Tap a hole.

A	B	
		Bench operations (con't)
		27. Cut threads with a die.
		28. Rivet.
		29. Remove a broken tap.
		30. Remove a broken stud.
		31. Dress a screw driver.
		32. Drive studs by hand.
		33. ______
		34. ______
		35. ______
		Drill press operations
		1. Grind a twist drill.
		2. Drill a hole to a layout.
		3. Drill a blind hole.
		4. Drill a hole with a jig.
		5. Set up and secure work on the table.
		6. Set up work in a vise.
		7. Drill cylindrical work.
		8. Drill holes at an angle with the surface.
		9. Ream a hole.
		10. Counterbore.
		11. Countersink.
		12. Spot-face.
		13. Tap a hole.
		14. Bore a hole.
		15. Bore a hole using buttons.
		16. Drive studs.
		17. Seat.

A B

Drill press operations (cont'd)

18. ____________________

19. ____________________

Engine lathe operations

1. Grind lathe tools.
2. True a live center. (soft)
3. Grind a hard center.
4. Adjust the tailstock to eliminate taper.
5. Mount a chuck on a lathe.
6. Mount a faceplate on a lathe.
7. Mount work on a faceplate.
8. Test for allignment of centers.
9. Drill a center on the drill press.
10. Drill a center with the drill in the headstock.
11. Drill a center with the drill in the tailstock.
12. Rough and finish turn work held on centers.
13. Face work held on centers.
14. Turn work to a shoulder.
15. Face a shoulder.
16. Knurl.
17. File.
18. Polish
19. Undercut work held on centers.
20. Turn work with a follower rest.
21. Set up work with a steady rest.
22. Turn a taper with a taper attachment.
23. Turn a taper or an angle with the compound rest.
24. Turn a taper by off-setting the tailstock.
25. Mount work in a three jaw universal chuck.

A	B	
		26. Mount work in a four jaw independent chuck.
		27. Center hollow ring work in a chuck.
		28. Face work in a chuck.
		29. Cut off work in a chuck.
		30. Drill.
		31. Bore.
		32. Ream with a machine reamer.
		33. Ream with a hand reamer.
		34. Turn irregular forms with the compound rest.
		35. Form with a forming tool.
		36. Bore a taper with the taper attachment.
		37. Bore a taper with the compound rest.
		38. Cut an American National Form Thread.
		39. Cut an Acme Thread.
		40. Cut a square thread.
		41. Cut a taper thread.
		42. Cut a left-hand thread.
		43. Cut a thread using a thread chasing dial.
		44. Cut an internal American National Form thread.
		45. Cut an internal Acme thread.
		46. Cut a multiple thread.
		47. Cut threads with a self-opening die.
		48. Measure threads with a thread micrometer.
		49. Measure threads by the three wire method.
		50. Tap threads with work held in a chuck.
		51. ________________________________

A	B

Engine lathe operations (con't)

52. ____________________

53. ____________________

54. ____________________

55. ____________________

Milling machine operations

1. Set up a cutter on the arbor.
2. Set up an end mill.
3. Set up a vise.
4. Indicate a vise jaw.
5. Set up work in a vise.
6. Set up work in a fixture.
7. Clamp work on the table.
8. Set up an index head for direct indexing.
9. Set up an index head for simple indexing.
10. Set up an index head for differential indexing.
11. Set up an index head for spiral milling.
12. Mill rectangular work.
13. Square the ends of rectangular work.
14. Straddle mill.
15. Mill a square or a hexagon.
16. Mill a spur gear.
17. Mill a bevel gear.
18. Mill a spiral gear.
19. Mill a spiral groove.
20. Mill the flutes of a reamer or similar tool.
21. Mill a worm gear.
22. Mill a screw thread.

A	B	
		Milling machine operations (con't)
		23. Mill a key way or similar groove.
		24. Mill a Whitney (Woodruff) keyseat.
		25. Saw with a slitting saw.
		26. Drill.
		27. Bore.
		28. Form mill.
		29. Mill a rack with a rack cutting attachment.
		30. Mill grooves with a slotting attachment.
		31. Mill with a rotary table.
		32. Mill with a vertical attachment.
		33. Mill special shapes with a fly-cutter.
		34. Profile.
		35. ____________________
		36. ____________________
		37. ____________________
		Shaper operations
		1. Set up work in a vise.
		2. Set up work in a jig.
		3. Set up work on the table.
		4. Set up work on an angle plate.
		5. Plane a horizontal surface.
		6. Plane a vertical surface.
		7. Plane an angular surface.
		8. Plane adjacent surfaces square with each other.
		9. Plane an end square.
		10. Plane an irregular surface.
		11. Plane a tongue.

A	B	Shaper operations (con't)
		12. Plane a groove.
		13. Plane a slot or keyway.
		14. Plane a keyway in a gear or pulley.
		15. Plane a T-slot.
		16. Plane gear teeth.
		17. ______________________________
		18. ______________________________
		Grinder operations
		1. Dress the face or a grinding wheel.
		2. Dress the side of a grinding wheel.
		3. Surface grind a flat surface.
		4. Face grind with the side of the wheel on a surface grinder.
		5. Grind an angular surface on a surface grinder.
		6. Form grind on a surface grinder.
		7. ______________________________
		8. ______________________________
		Miscellaneous operations
		1. Cut stock with a power hack saw.
		2. Anneal metal.
		3. Harden and temper tool steel.
		4. Case harden mild steel.
		5. Grind valves.
		6. Fit by lapping.
		7. Pack stuffing boxes.
		8. Lace leather belts with rawhide.
		9. Lace leather belts with metal lacing.
		10. Cement leather belts.
		11. ______________________________

APPENDIX H

LETTER THAT ACCOMPANIED CHECK LIST

193 Morningside Drive East
Bristol, Conn.
December 8, 1947

Mr. George M. Crook
Pine Manual Training School
Ansonia, Conn.

Dear Mr. Crook:

Inclosed with this letter you will find a brief bibliography and review of literature and two copies of a check list of machine shop operations. One copy of the check list is for your own use. The other copy I would like to have checked and returned to me as a part of our preparation for the performance test. Please initial or sign the last page of the copy that you return.

The check list was made up originally from:

1. Machine Tool Operation, Parts I and II by Henry D. Burghardt,

2. Improving Instruction in Industrial Arts by the A. V. A., and

3. Machine Shop Training Course, Parts I and II by Jones.

The original check list was cut down somewhat to include only those areas in which we are giving instruction.

From these operations we must select those that are to be included in the test problem. We should keep in mind certain principles when making our selections:

1. The operations should be as representative of the whole group as possible.

2. The operations must be such that they can be performed <u>safely</u> by a pupil who has had approximately 200 periods of instruction.

3. The operations must be those that can be objectively scored with equipment that we have in our shops.

4. The total group of operations must be such that it can be combined in a test problem that can be completed in four periods, or less, by the average pupil.

A performance test in shop work is always a compromise between the desire to include more operations in order to make the test more reliable and the necessity of keeping the test short enough that it can be administered in the average shop.

Will you indicate the operations that you think should be included in the test problem by placing a check mark in column A opposite each of those operations? Use your own judgement as to the number to include.

Space is provided to add operations.

I would appreciate having the check lists back before Christmas, if possible, so that I can work on the test problem during the vacation.

Cordially yours,

Richard B. Smith

APPENDIX I

INTERVIEW REPORT FORM

Machine Shop Performance test
Selection of Operations

(Supplementary interview report)

Date of interview___/___/___

With ____

At ____

- - - - - - - - - - - - - - - -

The first circulation of the Check List of Machine Shop Operations did not produce enough unanimity in the selection of operations. Six operations were checked by seven instructors, twenty-one operations by six instructors, and twenty operations by five instructors.

The purpose of this interview is to:

1. Find, if possible, the reason for the failure to obtain satisfactory results from the first checking of the list.
2. Obtain a smaller list of operations on which all can agree.

- - - - - - - - - - - - - - - -

Report of Interview

1. What was the "basis of checking" that you used when originally checking the Check List of Machine Shop Operations? (Check one)

   a. All the operations that you teach. ( )

   b. All the operations that you think should be in a senior high school machine shop course. ( )

   c. All the operations that you think are suitable to be included in a performance test. ( )

   d. A limited number of operations that you think can be combined in a test problem of appropriate length. ( )

   e. ____
   ____ ( )

2. How many periods of 45 minutes each do you think the performance test should require?

   (Circle one) 1 2 3 4 5 6 7 8 9 10

3. Approximately how many machining operations could such a test include?

(Circle one) 1 2 3 4 5 6 7 8 9 10 11 12

14 15 16 17 18 19 20

4. Review the Check List of Machine Shop Operations and select this number ( ) of machining operations that can be combined into a test problem. Include layout operations in this list. Note the following in each case:

a. The probable number of minutes the operation will consume. Write this in column A.
b. How the pupil will check the accuracy.

c. How the test administrator will check the finish.

5. Check the set-up operations that will be required for the machining operations you have listed. Write the approximate number of minutes that will be required for each set-up operation in column B.

APPENDIX J

LIST OF OPERATIONS SELECTED FOR THE TEST

Operations checked by six instructors in the second checking of the Check List of Machine Shop Operations. Listed by frequency of checking.

Checked by five instructors

Rough and finish turn work held on centers

Turn to a shoulder

Cut an American National Form thread

Checked by four instructors

Drill a center with the drill in the tailstock

Face a shoulder

Checked by three instructors

Lay out holes for drilling

Face work held on centers

Undercut work held on centers

Turn a taper or an angle with the compound rest

Set up a cutter on the arbor (Milling)

Set up work in a vise (Milling)

Checked by two instructors

Lay out centers in round stock with a center head and scriber

Tap a hole

Drill cylindrical work

Mount a faceplate on a lathe

Set up work in a four jaw chuck

Drill on a lathe

Bore

Ream with a machine reamer

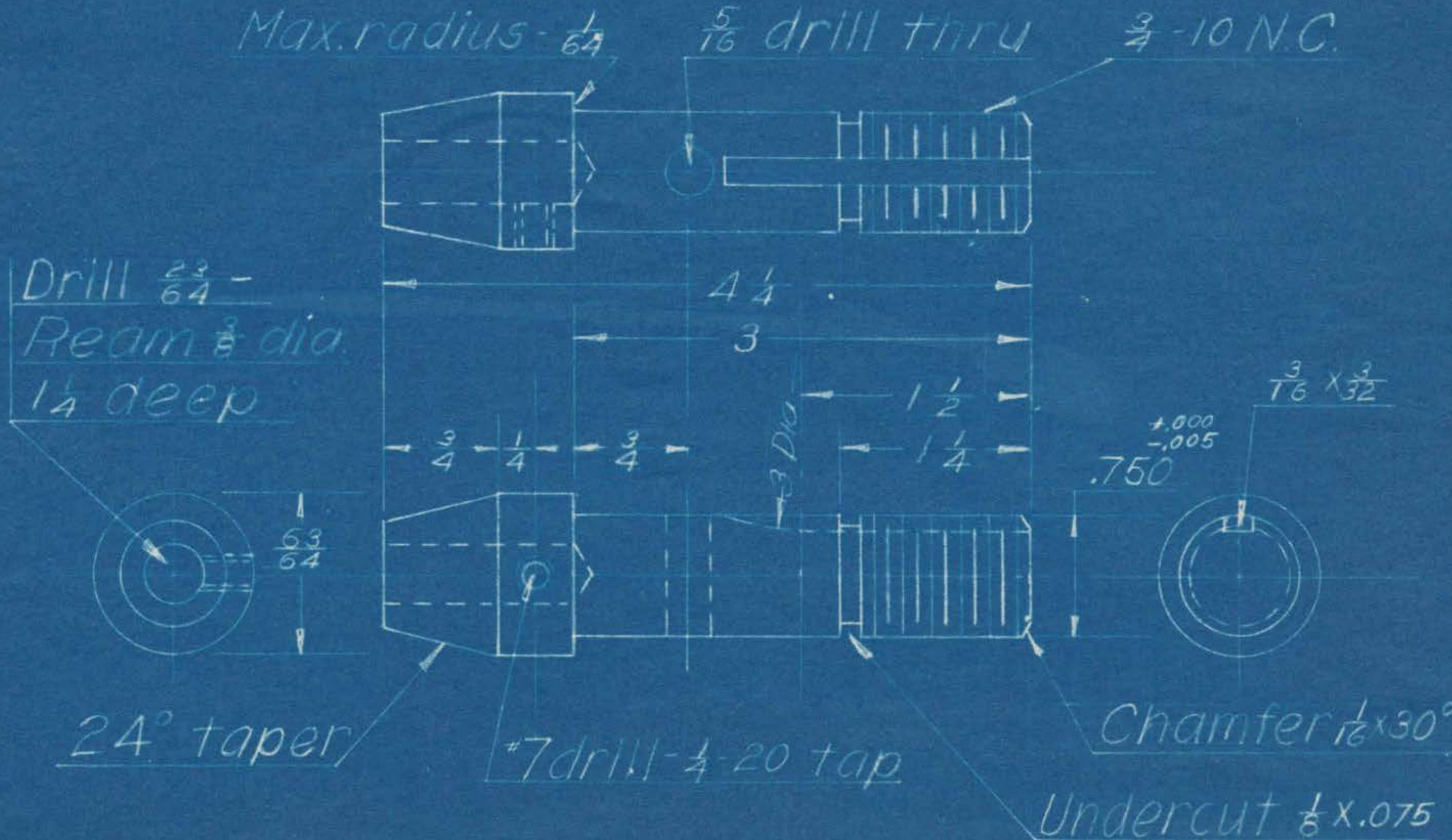
Mill a keyway or similar groove

Mill a flat on round stock

APPENDIX K

FIRST FORM OF THE BLUEPRINT

Performance Test Problem	Material CRS 1 dia × 4 $\frac{5}{16}$	All fractional tolerances ± $\frac{1}{64}$	Scale 1" = 1"



APPENDIX L

BLUEPRINT QUESTIONNAIRE

Machine Shop Performance Test
The Blueprint

Questionnaire. Made out by

1. Is the method of presentation of the material on the B/P satisfactory to you?____. If not indicate any suggested changes:

Item on the B/P	Suggested change
a.	
b.	
c.	
d.	

2. Are there any features in the test problem that you would eliminate because of equipment deficiencies? _____List them:

   a.

   b.

   c.

3. Are there any features in the test problem that you would eliminate because of safety?_____ List them:

   a.

   b.

4. Are there any features in the test problem that you would eliminate because of difficulty? ____List them:

   a.

   b.

   c.

5. Are there any features in the test problem that you would eliminate because of the length? _____List them:

   a.

   b.

   c.

APPENDIX M

EQUIPMENT AND TOOLS QUESTIONNAIRE

Machine Shop Performance Test
Equipment List

Questionnaire - Made out by

1. Can you meet the specifications for the machines?____If not, list any suggested changes:

   Lathes:

   a.

   b.

   c.

   Milling machines:

   a.

   b.

   Drill press:

   a.

   b.

2. Can you supply the accessories listed for each machine?_____For how many lathes?_____ For how many milling machines?_____ List the items that you are short on:

   a. b.

   c. d.

   e. f.

3. List here any of the small and hand tools that you do not have. Can you supply other sizes?

Item lacking	Suggested substitute
a.	
b.	
c.	
d.	
e.	
f.	

APPENDIX N

INSTRUCTIONS TO PUPILS QUESTIONNAIRE

AND INSTRUCTIONS TO PUPILS

Questionnaire. Made out by

1. Are there any parts of the instructions that are not clear?_____
   List:

   a.

   b.

   c.

2. Are there any points that might come up in the administration of the test that the instructions do not cover?_____List them:

   a.

   b.

   c.

   d.

   e.

   f.

3. Are there parts of the instruction that could not be applied in your shop?_____ List them:

   a.

   b.

   c.

   d.

   e.

   f.

Note; Mark or underline the attached copy if necessary to make your comments clear.

Machine Shop Performance Test
Instructions to Pupils

Note to administrator: These instructions are to be read as printed and are not to be enlarged upon or explained, but only repeated.

Instructions to pupils about preparation for the test which should be given an adequate time in advance of the actual date of the test:

> "On (Give date of test) you are to start on a performance test problem. In preparation for the test you are to see that your lathe tools are in good condition. You will be provided with thread cutting tools. You should also provide yourself with a 6" steel rule and a pencil."

Instructions to pupils at the start of the test; hand the pupils the blueprints and say:

> "You are to make this piece as a performance test. Examine your blueprint carefully. (Pause fifteen seconds) You may ask questions about the blueprint later if you wish to, but any answers that are given to you will be entered on your score sheet and your score will be reduced accordingly.
>
> "You will do your lathe work on these lathes: (Indicate a lathe for each pupil) These lathes are free from taper to a degree satisfactory for this job and have good centers. You will use the (Indicate which) milling machines for the milling operation. You will obtain your tools in the regular way. Undercutting and thread cutting tools will be provided.
>
> "You may check and oil your lathe before starting the test. Your time will start when you are handed your piece of stock. You may take as much time as you need for the test, but your speed will be a part of your final score. If you have to stop work at any time because of inability to obtain a proper tool or machine, inform me immediately so that the time so lost will not be counted as working time. Also you must inform me again when you are ready to resume work. If you have to continue the test on a later day, inform me when you actually stop work on the test, hand in your test piece and then clean up your machine. On the next class day your time will start when you have your work set up ready for machining. Inform me when you are ready.
>
> "The completed test will be scored on the basis of accuracy and finish as well as speed.
>
> "If you should spoil the job as far as one dimension or operation is concerned, proceed with the test and complete the rest of it according to the blueprint specifications. You will not be given a second piece of stock.
>
> "Do you have any questions?"

In case of questions, reread the part of the instructions that applies to the question asked.

Instructions to pupils when they restart the test on a subsequent day:

> "Set up your work in your machine just as it was when you stopped work the last day you worked on the test. Have me check it. When I tell you to start you may start working on the test."

APPENDIX O

QUESTIONNAIRE-CRITICISM -- 1948

MACHINE SHOP PERFORMANCE TEST

Questionnaire-Criticism

on experimental use

May-June, 1948

Filled in by ____________________

__________High School

I. The test problem blueprint

A. The operations

1. In the light of your experience with this test problem, do you think that there are operations in the test problem that should be omitted because of (Please check)

   a. Difficulty for the pupil___
   b. Unsatisfactory results in inspection___
   c. Unwieldy length of the problem___

2. Please list below the operations and state your reason for thinking that each one should be eliminated.

Operation	Reason for elimination

3. Would the test be more satisfactory from the point of view of administration and the results obtained if it was limited to lathe operations?_________

4. Do you suggest any changes in the tolerances allowed? ________List the changes here.

Old dimension and tolerance	New dimension and tolerance	Old dimension and tolerance	New dimension and tolerance

B. Dimensioning

1. Which of the following dimensions or specifications on the blueprint did the pupils show a tendency to miss entirely? Please check and list any others.

   a. 30° chamfer__ b. Undercut__ c. Depth of reamed hole__

   d. 90° relation of the 5/16" drilled hole and ¼-20 tapped hole__ e. #7 tap drill__ f. Break edges__

   g. Others____________________________________

2. What changes in the method of dimensioning would you suggest?

   a.____

   b.____

C. Questions

1. Do you think it is worth while trying to keep a record of the questions asked by the pupils about the blueprint?_____

2. Do you think that the pupils pick up any information that they need by observation of other pupils in a way that the instructor cannot prevent?_____

3. Were the questions that were asked due to (Please check)

   a. Inability to read the blueprint _____
   b. Lack of confidence _____
   c. Desire to stall _____
   d. Other cause (Please give other cause)____________

II. Machines and equipment

A. Which of the following difficulties did you encounter in maintaining machines and equipment as specified in the manual? Please check and list any others.

1. A 5/16" drill sharpened so that it would drill the proper size ___
2. A 3/8" reamer that would ream the proper size ___
3. Lathe chucks sufficiently accurate ___
4. Drill chucks sufficiently accurate ___
5. Proper run-out of the live centers ___
6. Proper taper of the lathes ___

7. ____

8. ____

B. Check any of the following suggested changes if you think that they would make the test more satisfactory. List any others that you suggest.

1. Supply all lathes with thread chasing dials ___
2. Make use of thread chasing dials optional ___
3. Substitute independent chucks for universal chucks ___
4. Have the milling machine set up with the vise on the table and indicated and the cutter on the arbor ___
5. Use a keyed instead of an unkeyed vise on the milling machine ___

C. How were small tools handled? Check more than one method if both apply.

1. In a tool crib with an attendant ___
2. In a tool crib without an attendant ___
3. On panels in the shop ___
4. In a special location for pupils taking the test ___

5. ____

III. Time record

A. On page eight of the manual you are asked to record the set-up time on the milling machine. Please check in the list that follows the steps that you included in that set-up time.

1. Mounting the arbor in the spindle of the milling machine ___

2. Mounting the cutter on the arbor ___
3. Mounting the vise on the table ___
4. Indicating the vise ___
5. Mounting the work in the vise ___
6. Centering the work under the cutter ___
7. Setting the depth of cut ___

IV. Inspection

A. Methods

1. List below any of the 29 inspection points in the Inspection Record that cannot be checked satisfactorily by the methods given in the manual.

a.______ b.______ c.______ d.______

e.______ f.______ g.______ h.______

2. 5/16" drilled hole

a. Is the method given for checking the centering of the 5/16" drilled hole satisfactory? ___
b. Would it be better to use a barrel or hole micrometer? ___
c. Do you have such a micrometer in your shop? ___

d. Do you suggest any other method? ________________

____

3. $\frac{1}{4}$ -20 tapped hole

a. There are three inspection points on this hole. Do you think this fact gives too much weight to the operations performed on this hole? ___

b. Which of the inspection points would you eliminate?

1) Depth ___ 2) Diameter___ 3)Squareness___

B. Check any of the following inspection points that you think should be included in the scoring of the test.

1. The alignment of the keyway in relation to the center-line of the work___
2. The centering of the keyway___
3. The diameter of the 5/16" drilled hole___
4. The 90° relationship of the 5/16" drilled hole and the ¼-20 tapped hole___
5. Concentricity of the reamed hole, taper, .750" diameter, 31/32" diameter, etc.___

C. List any other inspection points that you suggest.

1. ____
2. ____
3. ____

V. Attitudes of the pupils

A. Did the pupils display any of the following attitudes that might have affected their performance? Please check and add any others.

1. Indifference___
2. Over-emphasis on speed with a loss of accuracy and finish___
3. Over-anxiety___
4. ____
5. ____

Would you write on this page any other criticisms or suggestions that you have.

APPENDIX P

SUMMARY OF REPLIES TO QUESTIONNAIRE

The following is a resume of the answers given on the Questionnaire-Criticism on the Machine Shop Performance Test. Items where three or more instructors expressed opinions are included. Seven instructors checked the questions.

1. Four thought that some items might be eliminated from the test because of unsatisfactory results in inspection.

   a. Two of these thought that the 1/16" x 30° chamfer might be eliminated.

   b. Two of these thought that the tapped hole might be eliminated.

2. Four thought that the test would be more satisfactory from the standpoint of administration if it was limited to lathe operations.

3. Three thought that the test would be more satisfactory from the standpoint of results if it was limited to lathe operations.

4. Five found that pupils showed a tendency to miss the required depth of the reamed hole.

5. Four found that pupils showed a tendency to miss the 90° relationship of the 5/16" drilled hole and the ½-20 tapped hole.

6. Four found that pupils showed a tendency to miss the fact that they were required to break all sharp edges.

7. Seven thought that it was not worthwhile keeping a record of the questions asked by pupils about the blueprint.

8. Four found it difficult to supply universal lathe chucks of sufficient accuracy.

9. Three found it difficult to maintain the proper run-out of the live centers.

10. Four found it difficult to maintain the required taper of the lathes.

11. Four thought it desirable to make the use of thread chasing dials optional.

12. Three thought it would be better to have the milling machines set up with the vises indicated and the cutters on the arbors.

13. Four thought it would be better to use a keyed vise rather than an unkeyed vise.

14. Five thought that three inspection points on the ½-20 tapped hole gave too much weight to that element of the test.

    a. Three of these thought that the check on the diameter of the thread might be eliminated.

15. Four thought that the alignment of the keyway in relation to the center line of the work should be checked.

16. Four thought that the 90° relationship of the 5/16" drilled hole and the ½-20 tapped hole should be included in the inspection.

17. Four thought that concentricity should be included in the inspection.

18. Five thought that over-emphasis on speed had an adverse effect on accuracy and finish.

APPENDIX Q

FORM FOR CRITICISM OF

INSPECTION METHODS, 1949

Machine Shop Performance Test
Criticism of Instructions for the Inspection of Test Pieces

4¼ length

3 length

1¼ length of thread

3/4 to 5/16 drilled hole

¼ to ¼-20 tapped hole

31/32 diameter

11/16 diameter

3/32 depth of keyway

.188 width of keyway

.750 diameter

Thread Pitch Diameter

.075 depth of undercut

1¼ depth of reamed hole

3/8 diameter of reamed hole

12° angle, turned

30° angle of chamfer

90° relationship of the 5/16 drilled hole and the ¼-20 tapped hole

¼-20 thread, depth

¼-20 thread, squareness

1/64 maximum radius

Shoulder, squareness

Centering, 5/16 drilled hole

Centering, keyway

End, flatness of threaded

Burring all edges

Finish of turned surfaces

Finish of faced surfaces

APPENDIX R

EQUIVALENT HALVES OF THE TEST

Machine Shop Performance Test
Equivalent halves of the test
Results of three lists submitted

Three agreements

4¼	3
3/4	¼
12°	30°

- - - - - - - - - - - - - - - - - - -

Two agreements

1¼ thread	31/32
11/16	1/16
3/32	.075
.188	3/8 ream
.750	P. D.
¼-20 squareness	90°
1/64 radius	Shoulder-square
End-flat	Burring
Finish 31/32 diameter	Finish .750
Centering keyway	Centering hole

- - - - - - - - - - - - - - - - - - -

Items of no agreement, arranged in a suggested order of difficulty from easy to difficult:

¼-20 depth
Finish on shoulder
Finish on tapered end
Finish on threaded end
Finish on 12° angle
1¼ depth of reamed hole

Pairing of these six items resulting from conference

¼-20 depth of thread	Finish on shoulder
Finish on threaded end	Finish on tapered end
Finish on 12° angle	1¼ depth of reamed hole

APPENDIX S

ITEMS DELETED FROM THE TEST FOR THE SHORT FORMS

ITEMS DELETED FROM THE TEST RESULTS FOR THE SHORT FORMS

For the calculation of rank-difference correlations with grades

Biserial r	Clark's formula	Lindquist and Cook's formula
4¼	4¼	1¼ thread
1¼ thread	¼	¼
¼	1/16	31/32
31/32	¼-20 thread square	12°
1/16	Shoulder	¼-20 thread depth
12°	Centering hole	Centering hole

For the calculation of the reliability

4¼
1¼ thread
¼
31/32
1/16
12°
.750
¼-20 thread square

B I B L I O G R A P H Y

BIBLIOGRAPHY

1. American Vocational Association. Committee on Standards of Attainment in Industrial Arts Teaching. Improving instruction in industrial arts. Washington, American Vocational Association, Industrial Arts Division, 1946. 96 p.

2. Bawden, W. T. Army trade tests. Washington, Govt. Print. Off., 1919. 28 p. (U. S. Bureau of Education, Industrial education circular No. 4.)

3. Bingham, W. V. Aptitudes and aptitude testing; [1942 reprint] New York, Harper and Bros., 1942. 390 p.

4. Burghardt, Henry D. Machine tool operation. New York, McGraw-Hill Book Co., 1941. 2 v.

5. Chapman, J. Crosby Trade tests. New York, Henry Holt and Co., 1921. 435 p.

6. Clark, E. L. A method of evaluating the units of a test. Journal of Educational Psychology, 19:263-5, April 1928.

7. Davis, Frederick B. Notes on test construction: the reliability of analysis data. Journal of Educational Psychology, 37:385-90, October 1946.

8. Garrett, Henry E. Statistics in psychology and education. New York, Longmans, Green and Co., 1947. 465 p.

9. Greene, Harry A., Jorgensen, Albert N. and Gerberich, J. Raymond. Measurement and evaluation in the secondary school. New York, Longmans, Green and Co., 1943. 670 p.

10. Holsinger, James L. Performance test; basic electricity. Industrial Arts and Vocational Education, 36:206-8, May 1947.

11. Hunter, Robert S. A performance test involving accuracy, finish and time. Industrial Arts and Vocational Education, 31:13-14, January 1942.

12. Hunter, Robert S. Aptitude tests for the machine shop. Industrial Arts and Vocational Education, 34:58-63, February 1945.

13. Jones, Franklin D. Machine shop training course. New York, The Industrial Press, 1941. 2 v.

14. Lindquist, E. F. and Cook, Walter W. Experimental procedures in test evaluation. Journal of Experimental Education, 1:163-185, March 1933.

15. Machinery's handbook for machine shop and drafting-room; [11th edition] New York, The Industrial Press, 1943. 1815 p.

16. Newkirk, Louis V. and Greene, Harry A. Tests and measurements in industrial education. New York, John Wiley and Sons, 1935. 253 p.

17. Occupational counseling techniques, their development and application. New York, American Book Co., 1940. 273 p.

18. Orleans, Jacob S. Measurement in education. New York, Thomas Nelson and Sons, 1937. 461 p.

19. Pittsburgh. University. Vocational Teacher Training Staff. A report of the trade testing program at the University of Pittsburgh, 1932 to 1934. Pittsburgh, Pa., 1934. 50 p. mimeographed.

20. Richardson, M. W. and Kuder, G. F. The calculation of test reliability coefficients based on the method of rational equivalence. Journal of Educational Psychology, 30:681-87, December 1939.

21. Smith, Max. Relationship between item validity and test validity. Teachers College Record, 36:146-7, November 1934.

22. Toops, H. A. Trade tests in education. New York, Teachers College, Columbia University, 1921. 118 p. (Contributions to education No. 115.)

23. Wood, Ben D., Lindquist, E. F. and Anderson, H. R. Educational tests and their uses; Basic considerations. Review of Educational Research, 3:5-20, February 1933.