DESIGNS FOR SENSORY-MOTOR TESTS

AND

OTHER PSYCHOLOGICAL APPARATUS.

D.R. DE WET

Thesis submitted for the degree of Doctor of Philosophy in the Department of Psychology at the University of Cape Town.

1966

ACKNOWLEDGEMENTS

I am greatly indebted to my external supervisor Dr.S. Biesheuvel, who gave generously of his valuable time to the refereeing and critical discussion of my manuscripts, and made helpful suggestions for improvement.

Constructive criticism of the original draft, by Prof.K.Danziger, formerly my internal supervisor, was much appreciated.

Mr.A.O.H.Roberts, Mr.R.S.Hall, and Mr.J.M.Schepers were most helpful in elucidating queries on statistical theory and practice that arose from time to time.

Mrs.M.Barker typed the longer unpublished sections. The others were typed by Mrs.D.Griessel, Mrs.M.Steyn, Mrs.A.Gilbert and Miss S.Clelland. I wish to thank them all for their kind and efficient help.

Assistance received in test administration, computation, and other ways, is gratefully acknowledged in specific sections.

I must also thank Mr.A.Abramowitz, my present internal supervisor, for suggesting several improvements.

PREFACE

The design and construction of equipment for both basic psychological research and applied purposes in occupational selection and guidance, ergonomics, and training, is an important area that has hitherto received comparatively little recognition. Apparatus is often taken very much for granted as useful furniture affording some passive assistance in the study of behaviour, but nevertheless remaining something extraneous, that belongs more fittingly in the realm of mechanics.

Actually, the design and construction of testing equipment, no less than its application, is very intimately related to the measurement of psychological functions. Behavioural results can be no more sound than the instruments and techniques employed in deriving them.

This dissertation attempts to make some contributions of a twofold kind : Technological, comprising detailed illustrated descriptions of some original apparatus designs by the writer, and Behavioural, comprising accounts of research findings obtained with these, mainly on air-pilot candidates for the South African Air Force, and African industrial personnel.

A principal object in most of the designs, was the utilitarian one of producing valid instruments for the prediction of occupational suitability, but underlying this there was usually the intention that the instruments should also be adaptable for basic research purposes. While a few are rather specific to occupational selection for Aircraft-pilotage, Air-traffic control, and Mine winch-driving, the majority might be of wider interest.

Most of the apparatus consists of sensory-motor tests of various kinds, where the principal object is usually the observation and measurement of directed muscular movements. Of the others, three sets relate to visual-perception, where motor responses are generally of slight or even negligible importance; one is mainly concerned with the measurement of a temperamental function through the medium of visual-perception; and two are devices for administering cognitive tests.

Brief Summary of Contents

In the introductory chapter, the general significance of apparatus in psychology and the principles and practice of its development are discussed.

Section I is devoted to accounts of the technical construction and experimental applications of five tests of Co-ordination, a Steadiness test, and a complex Hand-foot Reaction test, designed primarily for measuring the sensory-motor skills of prospective air-pilots. All the tests were found to have useful reliability, and in all there were some measures that showed a significant relationship to success in flying training. Two of the tests that involved very fine patterns of muscular control, appeared to have the least usefulness for predicting success in flying. In two tests there was some evidence that the threat of physical punishment for incompetence, in the form of moderately unpleasant electric shock, improved the validity of a test in relation to a physically hazardous occupation like flying.

In Section II, technical and research details are given of thirteen Sensory-motor tests intended mainly for industrial research and classification. Three of these are preferably called "Assembly" tests, as they involve both sensory-motor skill and the ability to perceive three-dimensional form relations. Twelve of the tests showed useful reliability.

It was not possible to assess the reliability of one test as it was given as a continuous task once only. Two tests of fine Implement manipulation correlated significantly with a specific factory job. The one showed some connection with certain physiological and temperamental variates and also proved useful as a situation for observing temperamental behaviour. Results from the other indicated that (a) there were no significant differences in the respective performances of Africans and Europeans; (b) Good performers tended to estimate their achievement more realistically than inferior performers; (c) Most subjects were inclined to over-estimate what they could actually achieve, and the over-estimators tended to be less realistic.

Six simple skill tests, requiring moderately fine manipulation showed no validity for selecting operatives in a porcelain factory, which could partly have been due to the lack of specific and dependable criteria. However, useful information was gleaned on reliability, intercorrelations between tests, effects of educational level and mental ability, and differences in achievement attributable to differences in sex and age.

Three "Assembly" tests discriminated well between certain occupational categories of mine-workers, and showed substantial validity for winch-driver selection. Two rather specific Reaction tests also correlated significantly with winch-driving ability.

Diverse Sensory-motor activities, relationships between Sensory-motor tests, and between these and Mental tests, findings from various studies, the nature of the test sample, speed, tempo, accuracy, and other matters of practical or theoretical relevance in this field, are discussed in Section III.

Section IV deals with the construction and application of three sets of Visual-perception apparatus. The first, a Locus-estimation Test designed specifically for the selection of ground-control interceptor operators in the Air Force, showed useful predictive validity, and as the only practical test among others of the written sort, made a unique contribution to battery validity. The second, a Cancellation Timer, is a basic instrument intended for the administration of various printed Cancellation tests. While a test of Dots-cancellation applied by means of it to air-pilot candidates showed no validity as a measure of "sustained attention" in learning to fly, several findings of general interest emerged from the data, among which the most significant was evidence of parallelism in the performance curves of subjects who were initially good, median and poor. The third apparatus, a Dual-purpose "Fall" tachistoscope, is a basic instrument for presenting visual tests of (a) Speed of Perception, involving the identification of stimuli moving quickly across a restricted field, and (b) Span of Attention, involving the identification of stationary stimuli exposed briefly. A test of the former did not relate to achievement in flying, but a test of the latter correlated significantly with achievement.

Section V comprises the technical and research accounts of a device for obtaining measures of Flicker-Fusion threshold.

Although this function primarily involves visual-perception, there is evidence that it might have wider significance, particularly in the sphere of temperamental measurement. Very significant relationships were obtained between the critical fusion frequency (CFF) and the tendency to free expression of emotion, and also success or failure in flying training.

Two sets of apparatus, chiefly intended to provide controlled practical situations for measuring ability in various cognitive tasks of substitution or trial-and-error learning, are described in Section VI. Tests of paired-associate learning by trial-and-error, administered with these instruments, correlated significantly with certain flying-training criteria; a measure of Perseveration in a substitution task presented on one of the instruments, also showed significant validity. In addition to these results, information was obtained on rate of work and errors, recall, negative transfer, associative affinity, correlations with other variates, and the respective performances of English-speaking and Afrikaans-speaking groups.

In Section VII, certain principles, limitations and advantages pertaining to the use of practical tests in the vocational selection of personnel, are reviewed and evaluated.

CONTENTS

INTRODUCTION: Devising Psychological Apparatus.....1

SECTION I: SENSORY-MOTOR TESTS INTENDED MAINLY FOR AIR-PILOTS.

Introduction..........................19

Co-ordination and Floating-effect.............22

Handlebars: A Self-paced Test of Two-hand Co-ordination and some Results on Air-pilot Candidates...........33

A Variable Co-ordination Test and its Potentiality as a Gauge of Aptitude for Airmanship...............43

A Specific Arm-Leg Co-ordination Test for the Selection of Pupil Air-pilots..................57

The Roundabout: A Rotary Pursuit Test, and its Investigation on Prospective Air-pilots..............87

An Improved Steadiness Apparatus and its Validity for Air-pilot Selection..............102

A Portable Hand-foot Reaction Test..........119

Summary..........................133

SECTION II: SENSORY-MOTOR TESTS INTENDED MAINLY FOR INDUSTRIAL RESEARCH AND CLASSIFICATION.

Introduction..........................139

Two Tests of Implement Manipulation..........141
Simple Skill Tests Applied to Africans........169
Assembly and Sensory-motor Tests for African Mine Workers..........200
Summary....................241

SECTION III: AN APPRAISAL OF SENSORY-MOTOR ACTIVITIES.

Introduction..................247
Diverse Activities................248
Observation and Speculation..........251
Sensory-motor and Mental Abilities........253
Mental functions in Sensory-motor learning......257
Nature of the test sample............259
Grouping of Sensory-motor Abilities.....261
The work of Fleishman and Hempel..........269
Guilford's Scheme................274
Speed and Tempo..................276
Speed and Accuracy................281
Tables 1 to 6 283
Summary and Conclusions..............289
Abstract........................293

SECTION IV: VISUAL PERCEPTION APPARATUS.

Introduction.................... 304

A Locus-estimation Test............... 306

A Cancellation Timer.................. 319

Some Aspects of Performance in a Cancellation Test........... 322

A Dual-purpose "Fall" Tachistoscope........... 335

Measures of Speed of Perception and Span of Attention on Air-pilot Candidates............... 343

Summary........................ 356

SECTION V: AN APPARATUS OF OBJECTIVE TEMPERAMENTAL SIGNIFICANCE.

A Compact Flicker-fusion Machine and its Application to Air-pilot Candidates.......... 360

Summary and Conclusions 377

SECTION VI: APPARATUS FOR ADMINISTERING COGNITIVE TESTS.

Introduction.................... 380

A Multiple-choice Instructional Device....... 382

Paired-associate Learning through a Practical Medium.......... 399

A Serial Discrimeter for Studying Perseveration and Learning.......... 408

Summary........................ 436

Reprinted from "Psychologia Africana," Volume 10, 1963.

PSYCHOLOGIA AFRICANA, 1963, 10, 31–47

DEVISING PSYCHOLOGICAL APPARATUS

D. R. DE WET

Since the middle of the nineteenth century, when the application of experimental methods to the study of human behaviour began to rival the well-established mental philosophy, this empiric approach to psychological problems which had its inception as a confluence of philosophical speculation with the natural sciences, has come to depend more and more on specialized apparatus for the acquisition of its factual material. Indeed, it would be no exaggeration to say that without apparatus experimental and much applied psychology could hardly exist, for usually it is the only means whereby the investigator can control the main conditions under which his subjects react, and particularly in the study of individual differences, maintain a standard technique for all subjects.

Other advantages to be derived from the use of apparatus include the presentation of stimuli by the observer only when he is ready to record the subject's responses accurately; the repetition of experiments under the same conditions by either the same observer or others who wish to make an independent check on his findings; the systematic variation of certain conditions while others are held constant; and an objective measure of subjects' behaviour which is generally of greater scientific value than personal assessment by the observer.

This objective measure of response can be either qualitative, quantitative, or both. Quality is specific to a particular set of responses and there is no one technique suitable for measuring it in all. On the other hand, quantity of response lends itself to a standard procedure and can be measured either as the amount of response in a certain time, or as the time taken to produce a certain response. Time, variable or constant, affords a very versatile means of rating behaviour, and in psychological research, instruments for measuring time are among the most fundamental of all apparatus.

The earliest noteworthy applications of apparatus to the study of psychological problems were by physicists, astronomers, and physiologists, who very sensibly borrowed from the older sciences whatever laboratory equipment could be conveniently adapted to the new one. This was the start of the " brass-instrument psychology " of which William James disapproved so strongly. Influences of its distinctive apparatus-genre persisted with undiminished vigour until the late nineteen-thirties, and occasional models, archaic in appearance, but excellent in both workmanship and practical design, were still being produced by a few conservative makers even after the Second World War. Boring (4; p. 354) mentions that many older laboratories keep samples of Hering's colour-vision apparatus as show pieces, and that these " were beautiful specimens of German manufacture and played their part in earning for the new science the name ' Brass-instrument psychology '."

Psychology, like every science, is greatly dependent for its progress on developments in other sciences. Fechner's earlier researches on basic electricity (Brett, 6; p. 554) and likewise his knowledge of medicine and physiology must have been of value to him when in 1838–1840 he began to concentrate on experimental psychology. In 1838, Wheatstone's physical inventions and researches, and must therefore be made. Grings (15; p. 4) has aptly observed: " Because it will probably always be necessary to construct apparatus to fit specific research needs, the machine and electrical shops occupy positions of central importance in a psychological laboratory."

There is hardly any modern psychological activity in which apparatus of one kind or another does not play a significant part. Since Thorndike initiated quantitative experiments on animal learning in 1897 (Murphy, 20; p. 237), the demand for apparatus by animal psychologists has steadily increased, and although many basic recording and timing devices used on human subjects are also adaptable to animal research, there is still the continual necessity to design and set up equipment for specific research projects.

The social and ethnic departments of psychology, which have become so prominent during the last decade, particularly in Africa, are making considerable use of apparatus, mainly in the form of tests of individual differences suitable for group administration and inexpensive reproduction. The duplication of apparatus in quantity can be done most economically by commercial organizations equipped for this purpose, but before contracts for such work are assigned, it is usually necessary to evolve experimental prototypes in the laboratory.

A comparatively new area in which apparatus is proving its worth, is " engineering psychology," concerning the design of industrial and other equipment with a view to making it efficient and safe for human operatives. Actual experimental models provide the most fruitful method of pursuing what Fitts (12; p. 1287) calls " the study of man's behaviour in using the mechanical devices of our technological society."

Other notable fields are clinical work and personnel selection. Even in a collateral branch like statistics processing where most calculating equipment consists of industrially produced types used in all the mathematical sciences, there is also appreciable scope for the laboratory development of certain ancillary devices having a more direct bearing on the treatment of psychological data.

Apparatus serves as a powerful research tool, yet it would be unwise to assume that it provides a miraculous and infallible means of solving all the problems encountered by the psychologist. In common with the mechanical tools employed in its construction, it has limitations, and should be used for definite purposes and in the correct manner. Djinns of the calibre that emerged from Aladdin's lamp cannot be evoked by the manipulation of modern laboratory equipment, however cunningly contrived, and the research worker must simply resign himself to the fact that there is no *deus ex machina* capable of rendering him fantastic assistance.

Sources of information

From the purely practical standpoint, psychological apparatus can be grouped into two broad classes, namely, standard equipment that is commercially available, and that evolved to further particular investigations, and for which the call is too restricted to warrant production for the open market. Difficulty is seldom encountered in obtaining information about the first class as manufacturers issue catalogues, brochures, etc., covering things in current production. The second

Another type of work on applied physics, which is rather a treatise or discourse than a textbook in the accepted sense, yields fewer immediate returns to the searcher after practical ideas, but has considerable worth as " background " reading. An example is Whitehead (29) who discusses errors resulting from faulty elements in precision mechanism, and their relation to some general principles of design.

There are, then, numerous auxiliary ramifications of technical knowledge linked either directly or indirectly with the development of scientific apparatus, and Wilson (31; p. 70) describes the situation well when he says: " The born experimenter takes pleasure in learning his business and avidly picks up scraps of information about techniques, components, and methods wherever he can find them." Sometimes he may be faced with a problem so highly specialized that the relevant literature is either not easily obtainable or difficult to understand. The obvious alternative is to consult a person who is likely to know something about the matter—preferably an expert.

Apart from the transfer of specific knowledge that it effects, conversation can sometimes quite adventitiously give the designer a completely new perspective and sudden insight on other problems besides the one that has currently engrossed his attention. Expert opinion is not always imperative to induce this happy concatenation, and the ideas of non-technical people on certain aspects of design are by no means to be despised. As G. K. Chesterton has noted, " the last lesson the wise man learns is that the fool is sometimes right."

Design and constructional procedure

In evolving psychological apparatus, even in fields that have already been surveyed, it is hardly ever possible to start off by preparing sets of meticulous plans and specifications, although obviously the more intelligent foresight that can be brought to bear before construction starts, the better. James Watt, the distinguished Scottish inventor, is said to have maintained that a new device gets born on a workbench, not on a drawing-board, and this canny paradox will probably remain valid until man ceases to be a " tool-making animal."

Although there can be no complete lay-out in advance after the manner of an architectural drawing, it must not be assumed that the designer, at the first flash of inspiration, puts a piece of metal in the vice and attacks it furiously with a hack-saw. Before an attempt is made to give any part of the concept a three-dimensional form, a considerable amount of rough sketching and figuring usually takes place in a notebook, or on the backs of envelopes and match-boxes, margins of newspapers, table napkins, and other materials requisitioned for their ready availability rather than for any intrinsic virtue conducive to good calligraphy. The activity of brooding and gestation does not end when construction starts, but recurs frequently throughout the whole process whenever problems arise and modifications in the original scheme have to be introduced as possible ways of solving them. Meanwhile the designer's scrap-boxes acquire new contributions of discarded material (very likely thrown in to the accompaniment of maledictions) and his emotional tone fluctuates erratically between the nadir of dysphoria and the very pinnacle of optimism when another bright idea presents itself. This kind brass costs more than mild steel, but can generally be machined and finished more easily and is therefore preferable for many components. Contrary to popular belief, a hard, expensive wood often means less trouble to the workman than something softer and cheaper which *cuts* more easily, but is more prone to warp and split and may require special reinforcing, to say nothing of the difficulty in securing a presentable finish.

As research findings can really be no more meaningful or precise than the apparatus allows, a fair sense of proportion with regard to the necessary tolerances in workmanship should be maintained. Some measure of accuracy is always desirable, and certain things have to be very accurately made for proper functioning, but it is uneconomical to strive after high precision when a more moderate level will meet the requirements.

Often the successful operation of an apparatus will depend on the practicability of some central principle, and it is good policy to assemble and try out a rough elementary model of just this concept to determine its soundness before starting work on the rest. By this means the costly mistake can be avoided of doing much involved construction work on other parts that will prove worthless if the basic principle constituting the nucleus of the apparatus cannot be made to operate as planned. Having established that the main idea will work, it is fairly safe to proceed with the construction of the actual prototype, but in the case of complex apparatus there may still be considerable uncertainty about other points, and more detailed extensions of the rough-and-ready pilot model are to be recommended.

As the apparatus progresses, tests should be conducted at certain stages to ensure that the physical relationships are being kept in order, and that there is no marked deviation from the desired psychological objectives. Some apparatus cannot be tried out until it is pretty well complete. In other cases it may be possible to make trial applications of a sort during construction, and the sooner this can be done, even on a few subjects, the better, because the further the construction has advanced, the more difficult it becomes to carry out modifications.

Simplicity is an ideal to be sought in every design. Unnecessary complexity is a curse, partly due to muddled and immature thinking, and partly to conditioning in an age notable for the indiscriminate worship of "gadgets." As Grings (15; p. 3) says: "The more complicated the instrument becomes, the more subject it is to errors of various sorts; and the more complex it is, the more likely it is to be used incorrectly in a situation for which it was probably not designed." The smallest number of moving parts conformable with achieving the required function, and these of the plainest and most straightforward pattern, will greatly assist the construction process, make for consistent efficiency in operation and reduce the amount of irksome maintenance.

Another very desirable quality in any apparatus is that all the mechanism should be readily accessible for adjustments, cleaning, replacement of defective parts, or modification. Today, some industries are inclined to overdo the "sealed unit" type of construction, for three reasons. Firstly, this system is usually well suited to cheap mass-production; secondly, "style" and "sales appeal" must often have priority over utility; and thirdly, it is in the interests of manufacturers that the consumer should purchase an entire new unit rather than repair a defect and Wilson (31) treat it more concisely, and the latter (31; p. 107) sums the situation up when he says: " Mastery of the simple principles involved permits highly precise instruments to be built by an amateur."

Where mechanism in psychological apparatus has to be actuated, there are only four " prime movers " that merit consideration, namely, electricity, direct muscular power, spring clock-work, and gravity. Electricity is most convenient for laboratory work, and in other localities equipped with a permanent current supply, but it has disadvantages when this amenity is lacking and heavy batteries have to be transported and somehow kept charged.

Direct muscular power cannot be relied on when uniformity in the rate and quantity of the output is important.

The spring clock-work system, if properly constructed, can be very dependable and is better than the electric motor for apparatus that has to be really self-contained and conveniently usable in places remote from electric power supply. However, it will drive only comparatively light mechanism, and a continuous run on one full winding of the mainspring is of a certain limited duration.

Gravity is a source of power that is available to everybody, anywhere on earth, at all times, costs absolutely nothing, and for all practical purposes is wonderfully constant. The only creatures temporarily deprived of this universal force are passengers in capsules orbiting through space—but then it is accepted as an immutable natural law that those who would savour the pleasures of travel must endure certain inconveniences.

Gravity has been successfully applied to work a large variety of psychological apparatus, including fall-tachistoscopes and other exposure devices, chronoscopes, audiometers, and tests of visual co-incidence reaction, pursuit and other co-ordination. An outstanding specific example, which also conforms well to the ideal of simplicity in that it has few moving parts, is a method for measuring simple reaction time devised by Piéron (21) in collaboration with his assistant Boivin. A vertical metal rod in a stand falls freely when the observer releases a catch, and the subject has to check its fall by reacting on another catch. The distance of fall is read off on a scale, and his reaction time would be given by the formula $T = \sqrt{2D/G}$, where T is time in seconds, D is distance of fall in either centimetres or feet, and G is the acceleration due to gravity (980 cms. per second or 32 ft. per second).

In a simple and effective exposure apparatus which he called a " serial discrimeter," Seashore (24) used the principle of an elementary escapement system in conjunction with a weight and cord round a reel. The idea has many possibilities for application to other equipment.

A drawback of gravity-operated devices is that they sometimes have to be rather bulky, but this may well be outweighed by the advantages of cheapness, ease of construction and reliability in use.

Mechanical wear—harmful and helpful

The satisfactory functioning of a mechanism depends very much on the retention of proper constraint between moving or removable parts. When one surface rubs against another, there is an abrasive action that removes material and disc of glass is ground on another with a motion that is both to-and-fro and rotary, the lower stationary disc being worn convex and the upper disc, concave. The technique is well known to all amateur telescope builders, and described in Matthewson (18), Strong (26), and Wilkins and Moore (30).

These self-correcting methods of construction have the great advantage that the achievement of accuracy depends more on the process itself than on the proficiency of the workman.

Useful industrial stock

Of the many excellent commercial products having useful applications in the construction of psychological apparatus, and which it is generally either impossible or uneconomic to build in the laboratory workshop, the following are a few typical examples.

" Meccano " and similar construction kits on the principle of perforated metal strips, angles, and plates can be used for numerous experimental purposes and enable the designer to work in three dimensions at the very start of a project. Most " Meccano " parts are made to quite fine limits of accuracy, and some of them, particularly the gears, are solid enough to be incorporated in permanent apparatus.

" Handy-angle " is similar in principle to the " Meccano " angle-piece, but of thicker steel, much larger in cross-section ($1\frac{1}{2}'' \times 1\frac{1}{2}''$), and perforated with longitudinal slots. The makers also supply a standard gusset plate and bolts and nuts, which are very convenient. Structures from this material can be light, neat, and strong, and if increased rigidity is required, the joints may be welded. It has wide applications in the laboratory, and serves equally well for purely experimental models or final versions.

Another way of constructing frames and stands, etc., is on the tubular system, using pipe-stock and fittings in either steel or aluminium.

Clock-mountings, control-boxes, and bases for all sorts of both electrical and mechanical assemblies, can be made from blank steel radio chassis available in different sizes and formats; the use of such ready-made cases obviating the need for many hours of tedious sheet-metal work or cabinet-making.

Threaded brass or steel rod of various lengths and diameters to take standard nuts is sold by the bigger hardware stores. The quality is perfectly adequate for most purposes, and the cost almost negligible when compared with the bother of cutting long threads with hand dies or on a laboratory lathe.

Two types of industrial bearings most suited to psychological apparatus, where mechanism as a rule is not very heavy, are plummer-blocks and ball-races. The neatest and lightest plummer-blocks have bodies of die-cast aluminium and porous-bronze bushes. Self-aligning ball-races, although somewhat more expensive and bulky than the plain sort, present fewer difficulties in mounting.

Loose steel balls continue to be of value to experimental psychology as stimuli in tests of perception and sensory-motor capacity, and in other ways, such as the construction of special knobs, handles, and universal couplings. For the latter purposes it is necessary to fit stems or other fixtures rigidly to the balls, which can be done by soldering, brazing, or welding, but more satisfactorily by drilling practice here to enclose as much as possible, particularly movements like belts, gear-trains, and sprocket-drives, and furthermore, round off sharp corners and other projections that might inflict minor scratches and bruises, tear clothing, or even be just tactually uncomfortable and offensive to the aesthetic sensibilities. A point to note is that enclosure need not necessarily interfere with accessibility. By the judicious use of hinged doors or easily removable cowls or panels, the mechanism can be made readily accessible when required.

Discretion is always to be advocated in the use of sheet glass, which is a potentially dangerous material in any locality.

Handles and other controls to be operated by the observer should feel comfortable and work easily, and the same holds for those to be operated by the subject, unless the contrary is required for some definite research purpose. When it is not expressly intended to incorporate the influence of sheer physical fatigue, this element must be kept under proper control. The necessary output of energy in performing the task should be well within the power of the weakest subject.

If the observer's controls are few and simple, and all switches, knobs, etc., arranged in the logical sequence in which they will be used or according to some pattern of obvious relationships, there is less to distract attention from where it most properly belongs, namely, observing the subject's reactions.

The mere appearance of an apparatus can exert considerable influence on the subject's mental comfort, which is very noticeable in children, and to a lesser but still marked extent, even among sophisticated adults. The small, simple, and neat arrangement looks innocuous and encourages friendly co-operation; the large, complex, ugly machine appears ogre-like, or what Jack London would have called "pregnant with hurtfulness"—and repels. In the absence of a definite intention to promote unrest by the formidable appearance of the equipment, this *back action* effect, which at the best of times can never be a consistent, reproducible quantity, had better be kept subdued. Increased adaptation and familiarity during the progress of testing can be expected to reduce it, but not to a standard degree for all subjects. Concerning *back action,* Davis (10; p. 393) wrote: ". . . it is almost as though the experimenter is turning on a light to see what the darkness looks like. It must be the experimenter's goal to disturb the response as little as possible with his paraphernalia of measurement."

Apparatus tests of individual differences

The elaborate psycho-physical techniques of the early European psychology, characterized by an intensely detailed study of isolated phenomena on a few subjects, has been largely supplanted by the study of individual differences, originally founded by Galton. This involves measuring the traits of larger numbers of people and intercorrelating the results to determine the extent to which there is a general tendency for one trait to be associated with another.

The psychologist investigating individual differences is more concerned with relative than absolute measurement. Provided the units he uses remain constant for all the subjects in his investigation, it makes no difference what they are or how closely they approximate to any conventional scientific units. For example, if Pieron's (21) " Chronoptôscope " which was mentioned earlier is being used for

Most of the main principles and techniques in selection work concerning the development, validation and routine application of pencil-and-paper tests, are equally relevant to apparatus tests, and the literature devoted to these matters is among the most extensive in modern psychology. A concise, overall treatment is given by Bechtoldt (3).

Whether the test intended mainly for selection purposes will involve " apparatus " or not, the ideal should still be a twofold one of devising something which will give the maximum correlation of validity with the occupational criterion and also be reasonably identifiable in terms of some factor or set of factors; for in addition to the utilitarian purpose of evolving a reliable aptitude test to predict the candidate's suitability for certain aspects of a particular occupation, there is the important object of hoping to further the study of behaviour as an inductive science.

Conclusion

The design and construction of psychological apparatus can be systematized to some extent, but success also depends on ingredients that cannot be prepared by any stock formulae, such as the correctness of hypotheses in the light of known facts, and the evolution of the best compromise between the instrument and the end. The same would apply to the production of apparatus in any science.

When the designer feels himself handicapped and discouraged through his lack of proficiency in some specialized skill, there is consolation to be derived from a view expressed by Wilson (31; p. 120) which is typical of his acute practical insight: " Some experimenters succeed because of their magnificent technical skill, combined with adequate qualifications in other directions. Such skill is of very great importance, and is certainly worthy of cultivation, but in many cases its necessity can be avoided by intelligent design."

The ability to improvise—to make do with available resources and overcome difficulties *boeremanier* or pioneer fashion when conventional methods cannot be applied, is perhaps the most valuable single asset in apparatus design and, in fact, all research work.

Summary

1. In psychological research, apparatus is a useful but not a miraculous tool for controlling testing conditions and making objective measurements. There is the continual need to devise special equipment for use in all branches of modern psychology.
2. Some examples are given of the designer's main sources of information.
3. Practical procedure often involves trial-and-error, and always the necessity for compromise between the physical means and the psychological end. Other significant points are the ability to apply concepts that have proved successful elsewhere, the hypothesis, economy, accuracy, the pilot model and the final prototype, simplicity, accessibility and portability.
4. Elementary mechanical rules must be observed. Kinematic or geometric design can provide the solution to many problems. Sources of power for driving apparatus are compared.

23. REYBURN, H. A. *and* J. G. TAYLOR. On the interpretation of common factors: a criticism and a statement. *Psychometrika*, 1943, 8, 53–64.
24. SEASHORE, R. H. The Stanford motor skills unit. *Psychol. Monogr.*, 1928, 39, no. 2, 51–66 (Whole no. 178).
25. SEASHORE, R. H. Work and motor performance. *In:* Stevens, S. S., *ed. Handbook of Experimental Psychology*, pp. 1341–1362. New York, John Wiley, 1958.
26. STRONG, J. *et al. Procedures in Experimental Physics.* New York, Prentice-Hall, 1945.
27. TITCHENER, E. B. *Experimental Psychology; a Manual of Laboratory Practice.* Vol. 1, Qualitative Experiments; Vol. 2, Quantitative Experiments. London, Macmillan, 1924 and 1927.
28. WHIPPLE, G. M. *Manual of Mental and Physical Tests, 3rd ed.* Part 1, Simpler Processes; Part 2, Complex Processes. Baltimore, Warwick and York, 1921 and 1924.
29. WHITEHEAD, T. N. *The Design and Use of Instruments and Accurate Mechanism.* New York, Macmillan, 1934.
30. WILKINS, H. P. *and* P. MOORE. *Making and Using a Telescope, 2nd ed.* London, Eyre and Spottiswoode, 1957.
31. WILSON, E. B. *An Introduction to Scientific Research.* New York, McGraw-Hill, 1952.
32. WOODWORTH, R. S. *and* H. SCHLOSBERG. *Experimental Psychology.* New York, Henry Holt, 1954.

Mr. D. R. de Wet,
National Institute for Personnel Research,
P.O. Box 10319,
Johannesburg,
Republic of South Africa.

Manuscript received: 23rd November 1961.

CAPE & TRANSVAAL PRINTERS LIMITED, JOHANNESBURG

SECTION I

SENSORY-MOTOR TESTS INTENDED MAINLY

FOR AIR-PILOTS

SECTION I

Introduction

Air-pilotage is a vocation in which success depends not only on mental and temperamental attributes, but also, to a great extent, on practical skills, among which the following have come to be regarded as important :-

(a) Co-ordination : Directed muscular activity requiring the harmoniously integrated movements of several parts of the body; more specifically, the ability to use the hands, or the hands and feet, simultaneously, and without conflict.

(b) Distribution of Attention : The ability to make adjustive movements in response to several simultaneous stimuli, without becoming confused or concentrating unduly on some, so that others are neglected.

(c) Anticipation : The ability to judge the timing and amplitude of motor responses to allow for the inertia factor involved in objects that are moving.

(d) Smoothness of Movement : The performance of manual and/or pedal movements without roughness or jerkiness.

(e) Steadiness : Good muscular control, particularly under stress.

(f) Multiple Choice Reaction : Speed and efficiency in firstly, acquiring a series of sensory-motor habits, and secondly, applying them.

Five tests of Co-ordination, a Steadiness test, and a complex Hand-foot Reaction test, are dealt with in this section.

"Floating-effect", is a Two-hand Co-ordination test of the pursuit or automatically-paced type, specially designed to provide, in miniature, some of the genuine "feel" of actual air-craft control. In addition to Co-ordination, it is well loaded with Anticipation, Distribution of Attention, and Smoothness of Movement.

"Handlebars", a self-paced test of Two-hand Co-ordination, requires both co-ordination and very fine muscular control.

The Variable Co-ordination test, also self-paced, affords a large number of co-ordinative combinations between the limbs, but so far only the full combination of both hands and both feet has been investigated.

The Specific Arm-leg Co-ordination test for air-pilots, which had its origin in the Aptitude Tests Section of the S.A.A.F. during World War II, is of the pursuit type and simulates certain basic controls of an aircraft cock-pit.

The Roundabout, a complex pursuit test, was devised as a radical departure from the sedentary type of test usually applied to air-pilots. It involves co-ordination of most of the body musculature, and also Kinaesthetic factors, Distribution of Attention, Anticipation, Flexibility in reacting to sudden changes of "set", and Smoothness of Movement.

Steadiness of hand, under both normal and stressful conditions, is measured by an apparatus basically similar to an earlier war-time prototype, but of improved technical design.

In the Hand-foot Reaction test, which also has a war-time background, the candidate must first acquire the habits of moving the correct controls in response to certain signals, and then react quickly and accurately for a continuous period in accordance with these habits. The initial part of this test is mainly cognitive, as it requires trial-and-error learning.

Reprinted from the Journal of the National Institute for Personnel Research.

J. NAT. INST. PERSONNEL RES., 1959, 8, 28—38

CO-ORDINATION AND FLOATING EFFECT

D. R. DE WET

(Received 24th November, 1958)

During the Second World War, the increase of mechanisation in combat and industry caused emphasis to be put on the co-ordinative aspects of the motor-skills exacted from operatives.

This had particular reference to the more complex occupations fraught with hazard to valuable equipment and human life, of which aircraft-pilotage was (and still remains) an outstanding example.

Various tests of co-ordination for air-pilots were developed in the applied research units of most combatant countries, including the Aptitude Test Section of the South African Air Force under Lieut.-Col. Biesheuvel.

Some were based on existing tests of co-ordination like Moede's Two-Hand machine on the principle of the compound lathe-slide and the Koerth (1922) Pursuit-rotor; at least one, the "Carl Heinrich" Two-Hand machine used by the United States Army Air Force, was a hybrid combining the essential features of both the foregoing. Others had the addition of a pedal-response section and sometimes included supplementary distraction stimuli to further augment verisimilitude in the laboratory situation.

Complexity was carried to the extreme in one machine devised in the Aptitude Test Section of the S.A.A.F. Known as the Distribution of Attention Apparatus Mark IV, it incorporated three pedals, five hand controls and both visual and auditory distraction. The task was a continuous "pursuit" type and starting off with just one control it could be made more and more difficult by switching in successive combinations until the subject reached his limit of integrated co-ordination.

Most of these tests, notwithstanding the wide diversity of their individual design, have a common physical element which is one of several factors distinguishing the experimental set-up from actual flying conditions, viz., movement of the controls produces a direct and rigidly prescribed mechanical response by the machine.

In contrast to this, when an aircraft travels as an isolated solid entity through a fluid and elastic medium such as air, the transmission of control becomes "cushioned," there is greater latitude in mechanical response and some measure of inertia, which collectively evoke a very different pattern of potential kinaesthetic expectation and reaction on the subject's part. Thus the relationship between the aircraft in flight and its controls is not simple and direct like that between a motor-car and its steering wheel.

The greater sensitivity of aircraft response, coloured as it is by the peculiar nature of the "drift" or inertia in a "free-floating" body, demands some other kind of judgment to avoid over-control and over-correction.

Finally, there is the significant consideration that this shall function in more than one plane simultaneously.

As a broad "working" definition of all these cumulative influences, the term "Floating Effect" is suggested. From the behavioural standpoint it is largely absent in steering a land vehicle but prevails on one plane in vessels traversing the surface of the sea, becomes salient on two planes in submarines proceeding along below the surface, and assumes extreme prominence in the case of aircraft dependent on continuous motion for suspension in an attenuated fluid like air and having of necessity to be dirigible in three planes.

Although this phenomenon did not receive nearly the amount of attention it seems to have merited, nevertheless investigators were not entirely oblivious, as is evidenced by its inclusion in equipment such as the Link Trainer, the Rudder Machine of the A.T.S., and a "Triple Tester" designed by the Cambridge Psychological Laboratories.

The possibility of duplicating the actual flying situation to the extent of propelling the subject at speed and in true aerial suspension within the confines of an average room by some expedient like a miniature helicopter, is beset with insurmountable practical difficulties. But this limitation is not without advantage, for the copying of actuality in the test situation, when carried beyond certain bounds, may stultify a major purpose of the research which seeks to identify the *factors* involved in flying. If the test is made nearly as complex as the "real" thing, a similar obscurity may tend to persist over both.

One apparatus which attempted a compromise between "test" and "reality" was a Link Trainer fitted with special pendulum recording mechanisms and used in the South African A.T.S. during the latter part of the last war.

As the actuation principle used on the Link, as well as its control transmission, is almost entirely through the medium of compressed air expanding in a system of bellows, a good proportion of the "Floating Effect" experienced in actual flight was successfully reproduced. The main disadvantages of the machine were the excessive amount of space it took up permanently and the fact that its close approach to "reality" in other complex detail and "gestalt" mode of operation detracted from its usefulness as an analytical instrument.

To supply the need for a more controllable test situation of this "drift" or "Floating Effect," a device called the Rudder Machine was developed in the A.T.S. It utilised the principle of a swinging beam. Equilibrium was disturbed continuously by a motor-driven weight and counterpoise had to be accomplished by another moving weight coupled with wires to a pair of pedals operated by the subject who did not have the beam itself in his field of view, but interpreted its movements in terms of a single spot of light drifting across a large screen.

The Rudder Machine had three disadvantages:

(1) It was cumbersome and took up a great deal of floor space.

(2) The activity was restricted to one plane only.

(3) Excessive friction and variability could never be eliminated from the cable mechanism of the compensating weight.

A most ingenious little Pursuitmeter, known as a "Triple Tester," was produced in the Cambridge Psychological Laboratories during the last war and Prof. Sir Frederic Bartlett very kindly presented a specimen model to the South African A.T.S.

This device secured "Floating Effect" (in one plane only) by means of a mechanical gear transmission and a variable friction-clutch between the subject's control-wheel and the cursor which had to be steered to follow a track of dots on a rotating drum. The cursor was motivated through the transmission by a separate electric motor. Control of speed and direction by the subject was through the friction-clutch only.

The resultant "Floating Effect" proved very successful, but a serious drawback was the play and backlash in the complex transmission. Allowing, of course, for the model at that time being still to some extent in the prototype stage, a lot of this would no doubt have been overcome later by more rugged and accurate construction, but a completely reliable version of such mechanism would entail much slow and expensive precision work. Furthermore a separate clutch system would be necessary for any additional plane of operation.

The Floating Effect Test.

In the light of the foregoing, the apparatus to be described here attempts to provide a compact and moderately difficult co-ordination activity stressing the kind of judgment required in controlling and compensating sensitive movements in a task where a major consideration is the subject's allowance for those physical factors (in totality to be called "Floating Effect") which interpose between himself as the gubernative organism and his direct attainment of some definite psychomotor ideal.

Although designed mainly with a view to air-pilot selection, the principle may have a wider applicability to other occupations. If the "Two-hand" form is thought too limited for pilot testing or other purposes, then a modified machine having both hand and foot sections and mounted on an upright floor-stand should be quite feasible.

Mechanical Principle.

The "Floating Effect" test (Plates 1 and 2) distributes activity on two planes in a simple mechanical manner, moving parts being rugged and restricted to the bare minimum.
All main bearings (Plate 2, Figures 4 and

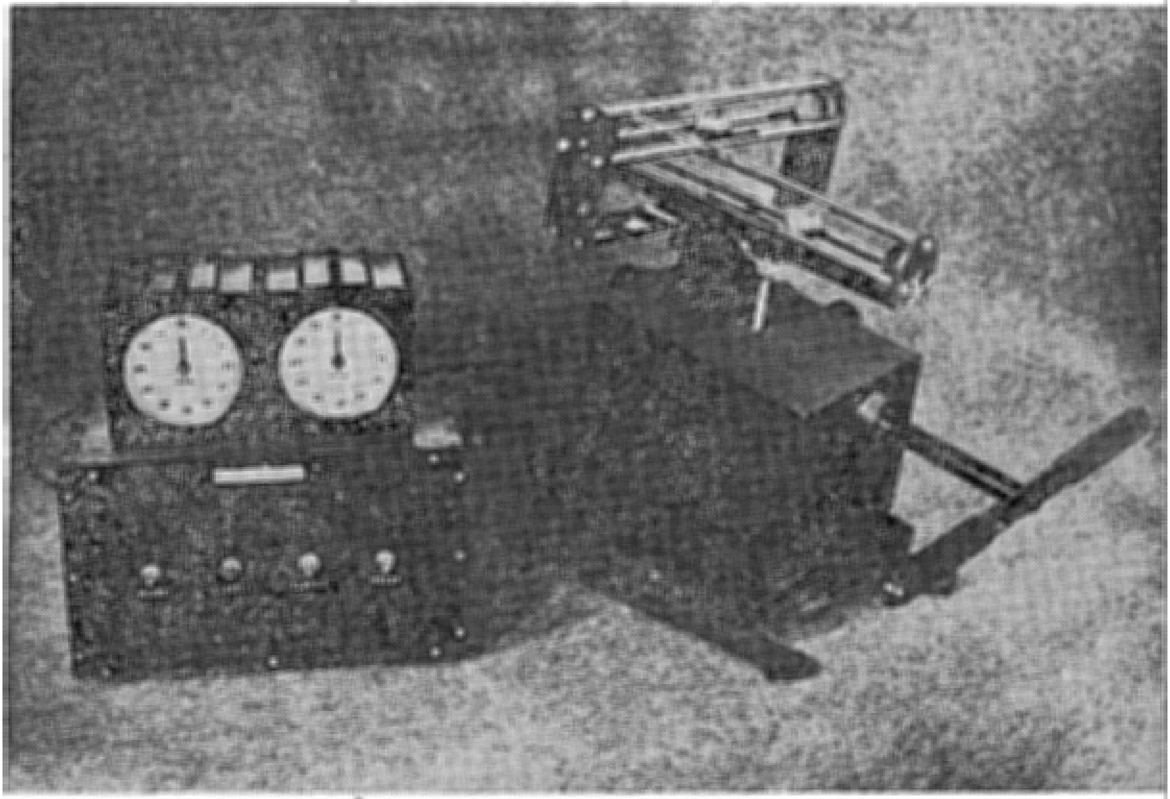


Plate 1: The "Floating Effect" Test

7) are such that the thrust is downwards so that wear will not produce "lost-motion" in operation.

The stimuli consist of two large steel balls running on respective sets of bars at right-angles to each other. (Plate 2, Figures 1, 2 and 8.)

This orthogonal framework pivots on a gymbal and through a rod and universal joint (Plate 2, Figure 9) is motivated by a constant-speed electric motor at two revolutions per minute to describe a steady "swash-plate" movement.

The structure of motor, housing and framework is mounted in an outer set of large gymbals.

By manipulating a pair of handles (Plate 2, Figures 1, 2 and 5) attached to the motor housing, the subject has to keep the framework level in two planes against the action of the motor which causes a continuous disturbance of the level.

The criterion of "levelness" is that the steel balls should always be balanced at the middle sections of their respective sets of bars where there are strips of insulating material. (Plate 2, Figures 1, 2 and 8.)

Electrical Circuit (Plates 1 and 3).

As a matter of convenience, the control panel is a separate table-unit (Plate 1). Recording is done on two "Standard" electric clocks (minutes and seconds) having clutch systems of suitably low voltage for exposed circuits (Plate 3).

The balls, in conjunction with their sets of bars, act as simple switches in series with the clock clutches, and continuity is broken while they remain on the insulated sections. Both clocks record total time off the criterion in a run of standard duration, one through the fore-and-aft bars, and the other through the lateral bars.

The driving motor is a "Klaxon" Type EM5CBI-M21, continuous rating with permanent capacitor.

25

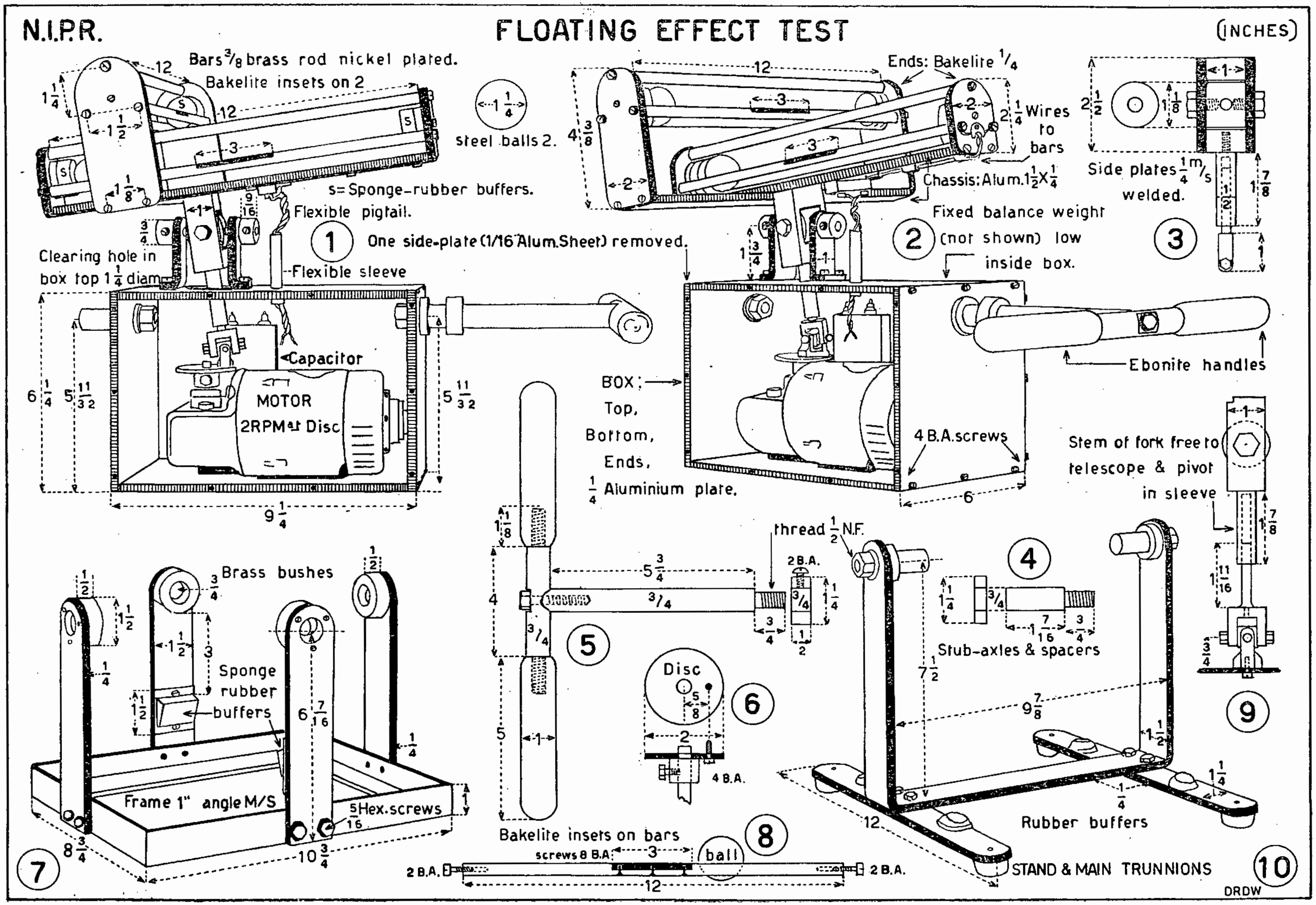

N.I.P.R.
FLOATING EFFECT TEST
(INCHES)
Bars 3/8 brass rod nickel plated.
Bakelite insets on 2
steel balls 2.
s=Sponge-rubber buffers.
Flexible pigtail.
One side-plate (1/16 Alum.Sheet) removed.
Clearing hole in box top 1 1/4 diam.
Flexible sleeve
Capacitor
MOTOR
2RPM at Disc
Ends: Bakelite 1/4
Wires to bars
Chassis: Alum. 1 1/2 X 1/4
Fixed balance weight (not shown) low inside box.
Side plates 1/4 m/s welded.
BOX: Top, Bottom, Ends, 1/4 Aluminium plate.
4 B.A. screws
Ebonite handles
Stem of fork free to telescope & pivot in sleeve
Brass bushes
Sponge rubber buffers
Frame 1" angle M/S
5/16 Hex. screws
thread 1/2 N.F.
2 B.A.
Disc
4 B.A.
Stub-axles & spacers
Bakelite insets on bars
screws 8 B.A.
ball
2 B.A.
Rubber buffers
STAND & MAIN TRUNNIONS
ORDW
1
2
3
4
5
6
7
8
9
10

Plate 2

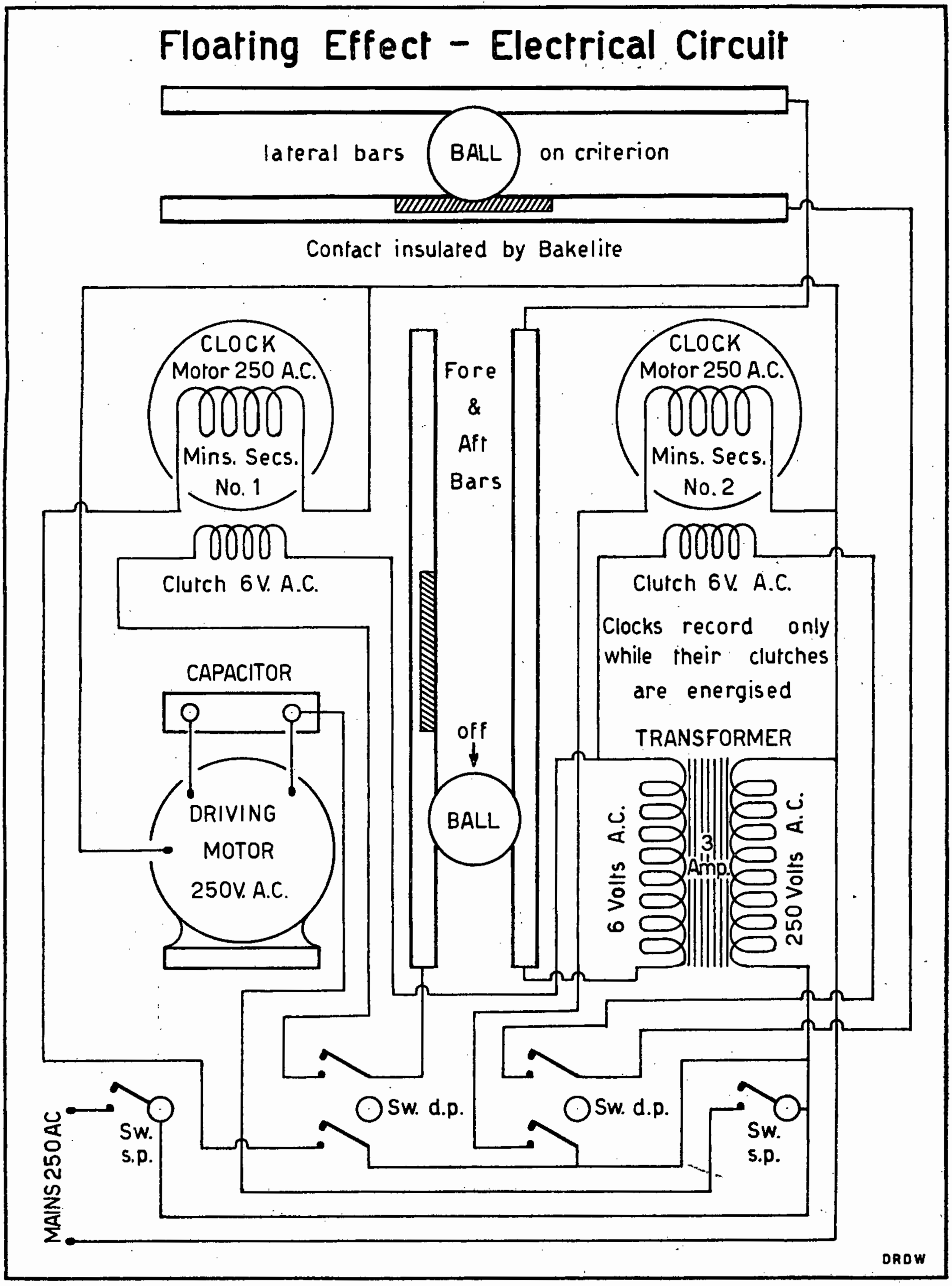

Floating Effect - Electrical Circuit
lateral bars
BALL
on criterion
Contact insulated by Bakelite
CLOCK
Motor 250 A.C.
Mins. Secs.
No. 1
Clutch 6V. A.C.
CAPACITOR
DRIVING
MOTOR
250V. A.C.
Fore
&
Aft
Bars
off
BALL
CLOCK
Motor 250 A.C.
Mins. Secs.
No. 2
Clutch 6V. A.C.
Clocks record only while their clutches are energised
TRANSFORMER
6 Volts A.C.
3 Amp.
250 Volts A.C.
MAINS 250AC
Sw. s.p.
Sw. d.p.
Sw. d.p.
Sw. s.p.
DRDW

Plate 3

Technique.

On this test subjects work standing. The machine rests on a table so that its handles project over the edge and the control panel is placed well to one side.

A standard starting position is kept for each trial: Left wing down and tail down, the handles being straight and "neutral."

Instructions to the subject are as follows:

"This is a test of your ability to control and compensate sensitive movements in two planes.

"Using these handles, you have to keep the framework level, so that the steel balls remain on the black strips at the middle parts of their tracks.

"Whenever you allow the steel balls to run off the black strips, error will be recorded against you. So when they run off bring them on again without delay.

"Concentrate on keeping *both* balls on the black strips by watching the framework to see in which direction it is tending to tilt so that you can anticipate the correct control to keep it level.

"You have three trials of three minutes each. When I give the word to start, bring the balls on to the black strips as soon as you can. Ready? Go!"

Experimental Application.

The test was administered to pupil-pilot candidates for the South African Air Force, the main objects of the investigation being as follows:

(1) To get some indication of "Floating Effect" test reliability.

(2) To obtain a tentative validation with the several Flying Training Criteria.

(3) To sound its potentiality as a predictor simply of success or failure on a course of Flying Training.

Follow-up on the training criteria was possible on 25 subjects who qualified in a 1952 group and 33 in a 1955 group.

In the former there was no available data on Formation Flying and in the latter none on Number of hours dual flying necessary in each case before the pupil was permitted his first solo flight.

The various criteria are based on instructors' reports containing their percentage estimates of pupils' performance in the several branches during training and on placement in the final "Wings Test."

"General Flying" is a unitary assessment of the pupil's all-round capabilities before he embarks on the "Wings Test" and can be regarded as a close pointer to his final performance.

It happens almost invariably that certain pupils are suspended from the course. This may be done for several reasons, such as medical unfitness which has developed during training, at the pupil's own request, because of inability to master the ground subjects, or most likely on account of some persistent deficiency under actual flying conditions which in the instructors' opinion precludes him from ever making an acceptable pilot.

This latter class of suspensions especially, affords a useful additional criterion for assessing the value of the test as a predictor of successes and failures by the biserial correlation method.

Suspensions for flying incompetence were 4 in 1952 and 11 in 1955, thus for the biserials the groups are 29 and 44 respectively.

Both groups were originally sections of larger samples (Table 1) who did a number of standardised tests both "intellectual" and sensory-motor on which recommendations for acceptance or rejection for flying training were based. "Floating Effect" played no part in this selection and was included in the battery only for experimental purposes. Nevertheless, since it must be assumed that both the standard test battery and the additional hurdle of a selection board have some validity, it is apparent that the validation of the "Floating Effect" test must be viewed against a backdrop of curtailment.

Table I

DERIVATION OF SAMPLES

	1952	1955
Applied for flying course	67	169
Accepted for training	30	51
Suspended from training (all reasons)	4	17
Suspended for poor flying ability	4	11
Completed course successfully	26	34
Number of latter who did the Floating Effect test	25	33

Results and Discussion.

It was unfortunately not advisable to pool the two groups as their respective means and S.D's (Table II) were significantly different.

The more heterogeneous 1952 group with the large S.D. showed a superior mean performance and the total of 29 subjects with an average age of 17.74 years ranging from 16 to 19 is made up as follows:

In addition to matriculation

Spare-time work with a radio firm	1
Radio transmitting experience	1
Electrical apprenticeship	1
Technical (N.T.C. II) qualifications	1
One year at university	1
Clerical work	2

Instead of matriculation

Technical (N.T.C. III) qualifications	1
Matriculation only	21

The more homogeneous 1955 group with the small S.D. showed an inferior mean performance and the total of 44 subjects with an average age of 18.27 years ranging from 17 to 21 is made up as follows:

In addition to matriculation

Technical (N.T.C. I) qualifications and a radio diploma	1
Apprenticeship in Aero Design	1
Clerical work	5
Matriculation only	37

On these small samples there was no indication that technical, clerical or post-matriculation experience as a whole influenced performance on the "Floating Effect" test.

"Technical" subjects were below the group mean in 1952 and above it in 1955. "Clerical" subjects tended to be average within their groups.

The largest proportion of subjects in both groups had no post-matriculation experience and the respective means of these two sections came to 54 for 1952 and 73 for 1955, which do not differ significantly from the total group means in Table II.

Table II

MEAN FLOATING EFFECT ERROR SCORES FOR SUBJECTS IN EACH GROUP

(Including those later suspended for flying incompetence.)

Floating Effect	1952 Group	1955 Group
Mean (Error)	55	72
N.	29	44
S.D.	15.878	10.436

The difference between the two means is significant at .1%.

The F-test shows that the S.D. difference is significant at 1%.

It is unlikely that the significant difference between the means in Table II is due to artifacts such as wear or maladjustment in the apparatus which could have made the test activity more difficult for the latter group. Any variation caused by slight spark-corrosion of the balls and bars through previous use, would on the contrary have made the task spuriously easier for the 1955 group by reducing the sensitivity of error recording.

A lower average age may have had some bearing on the superior performance of the 1952 group.

The proportion of suspensions (for all reasons) from Flying Training was 13.3% in 1952 and 33.3% in 1955. By the Chi-squares test this difference is significant at the 8% level, which suggests some flexibility in the criterion. It is of interest that the group with the high percentage suspension had the significantly inferior mean performance on "Floating Effect."

In Tables III and IV giving the Product Moment correlations between "Floating Effect" trials separate and combined and the Flying Criteria, the samples are of necessity small and the results should not be regarded as more than indications of general tendency.

Correlations are probably depressed to some extent through the effect of training that has smoothed away a certain amount of individual difference where the test might otherwise have shown its discrimination.

Table III
PRODUCT-MOMENT CORRELATIONS BETWEEN TEST TRIALS AND CRITERION
(Decimal points omitted.)
1952 GROUP. 25 SUBJECTS.

Floating Effect	Trial 1	Trial 2	Trial 3	Trials 1, 2, 3	Worst of 1, 2, 3
Wings Test	44	21	27	36	39
General Flying	36	22	38	34	30
Night Flying	27	20	48	29	27
Pilot Navigation	07	02	32	08	09
Hours Dual Before Solo	36	21	35	32	34
Instrument Flying	30	17	35	28	35

Less than 32 is below the 10% significance level.
38 to 45 falls between the 5% and 2% significance levels.

Table IV
PRODUCT-MOMENT CORRELATIONS BETWEEN TEST TRIALS AND CRITERION
(Decimal points omitted.)
1955 GROUP. 33 SUBJECTS.

Floating Effect	Trial 1	Trial 2	Trial 3	Trials 1, 2, 3	Worst of 1, 2, 3
Wings Test	26	18	15	29	30
General Flying	32	16	03	20	34
Night Flying	11	02	−10	05	13
Pilot Navigation	28	30	28	31	31
Formation Flying	30	02	06	12	25
Instrument Flying	26	28	13	25	28

Less than 28 is below the 10% significance level.
More than 32 is significant at the 5% level.

The "Worst of three trials" was included on the hypothesis that a measure of the subject's lowest limit of proficiency might be important in relation to his performance in flying training. But the present samples are too small to warrant any conclusions on this point. There appears no significant difference between the "Worst of three trials" and the first trial because the first is usually (though not invariably) the worst.

With the exception of Pilot Navigation, the correlations for the more heterogeneous group in Table III show, as could be expected, a tendency to be larger than comparable ones for the more homogeneous group in Table IV.

On four of the five variables and Pilot Navigation there is a reversal of this expected pattern, and although none of the differences between corresponding items reaches the 15% level of significance, nevertheless their recurrence along four cells in the row looks too consistently anomalous for chance fluctuations. If it is permissable to treat this as a case of compound probability, then the significance approaches the 10% level.

A careful re-check of the raw data revealed no errors in coding.

The correlations between Trial three and Night Flying (.48 in Table III and —.10 in Table IV) are significantly different at the 2% level. It is possible that supervenient to any compositional disparity between the two groups, a change in flying training procedure had some influence.

The 1952 group did initial training on Tiger Moths before proceeding to Harvards, whereas the 1955 group began directly on Harvards.

Progress being more leisurely and less exacting on the former slow aircraft, definite sets of habits (both useful and otherwise) could have been engendered which persisted to some extent when pupils graduated to the faster Harvard machines. The higher correlation for the Tiger Moth group might be partially accounted for in terms of different orientation. The slow plane afforded more scope for flying by kinaesthetic and extraneous visual reference and more time for the correction of wrong direction and bearings.

In this respect the situation seems a nearer approach to the "Floating Effect" test.

The 1955 group began flying directly on the faster Harvards and under these circumstances Night Flying could conceivably have necessitated a greater insistence on theoretical navigational aspects and less reliance on "feel" and environmental clues for correct course and position.

As a "test—re-test" was not practicable, the inter-trial correlations in Table V must suffice as a measure of test reliability.

Table V

PRODUCT-MOMENT CORRELATIONS BETWEEN THREE SUCCESSIVE TEST TRIALS

Floating Effect	1952 Group. N. = 29		1955 Group. N = 44	
Trials	2	3	2	3
1	.81	.84	.84	.72
2	—	.83	—	.81

Average correlation is .81 which is significant at .1%.

Taken on three trials together the Reliability Coefficients come to .94 for the 1952 group and .92 for the 1955 group and the Standard Errors of Measurement to 4 and 3 respectively.

In Table VI the Biserial correlations between the test trials and the "Success-

Table VI

BISERIAL CORRELATIONS ON THE "SUCCESS-SUSPENSION" CRITERION
(Suspensions through lack of competence in actual flying.)

Floating Effect	1952 Group. N = 29			1955 Group. N = 44		
Trials	r bis	S.E.	r bis / S.E.	r bis	S.E.	r bis / S.E.
1	.83	.17	4.9	.24	.20	1.2
2	.63	.22	2.9	.35	.19	1.8
3	.66	.21	3.1	.25	.20	1.3
1 + 2 + 3	.71	.19	3.7	.31	.19	1.6

Suspension" criterion conform to expectation by being greater for the more heterogeneous group. Trial I is prominent here and also in Table III on the "Wings Test," the Main Flying Criterion.

An hypothesis that the initial stage in a sensory-motor learning series has importance in occupational forecasting would be supported by the findings of Woodrow (1956) on data from a Two-Hand Co-ordination test of the Moede type. She found that the first of five trials on this test had "particular value" as a predictor of flying-course suspensions.

A study of instructors' reports on the suspended pupils in both groups gave the general distribution in Table VII. The condensation of the main reasons for flying incompetence entailed some latitude in the qualitative interpretation of certain points mentioned by instructors and the arrangement adopted in this table lays no claim to rigid objectivity.

Lack of "Balance" is stressed at least once in each of five cases, and in one of these, twice. Lack of "Air Sense" is stressed in three cases, but in one of these there is slight preponderance in favour of lack of "Balance".

Table VII

GENERAL CLASSIFICATION OF ALL THE SUSPENDED PUPILS IN BOTH GROUPS BASED ON FLYING INSTRUCTORS' REPORTS

Main Reasons for Suspension	1952	1955	Total
Medically unfit; At own request, etc.	0	3	3
Failure in Ground Subjects	0	3	3
Poor Distribution of Attention	0	1	1
Poor Learning Ability	1	1	2
Nervousness; Lack of confidence, etc.	1	2	3
Slow reaction; No Anticipation or "Feel"	0	2	2
Lack of "Balance" and "Air Sense"	2	5	7

Biesheuvel (1942) gives a broad definition of "Air Sense" as: "An assessment of movement and position with regard to the new visual clues, particularly an ability to think in terms of three dimensions."

In the absence of more specific knowledge about this concept, it does not appear unreasonable to associate it behaviourally with "Balance."

Summary.

(1) A comprehensive description is given of a sensory-motor test that attempts to supplement co-ordination in the usual sense by making provision for such elements as kinaesthetic sensitivity, anticipation and spatio-temporal judgment.

(2) Test reliability as indicated by inter-trial correlations on two groups separated by a three-year period appears satisfactory.

(3) Correlations with the Flying Training Criteria were all positive, except one and with a few exceptions larger for the more heterogeneous group, where furthermore, a larger selection ratio implies comparatively less depression of the correlations through curtailment.

(4) Biserial correlations on the "Success-Suspension" Criterion were all posi-

tive and here again larger for the more heterogeneous group.

(5) In this latter group, the first trial appeared more prominent than the second or third trials both in relation to achievement on the Main Flying Criterion (Wings Test) and Suspension for poor flying ability.

Acknowledgments.

Mr. J. B. Kirstein took the photograph.

The test was administered by Mrs. V. R. Wilson and other members of the N.I.P.R. testing staff.

Mr. J. Tromp willingly and informatively discussed points relating to flying training procedure and the criteria.

The helpful interest shown by Dr. H. L. Watts with regard to a certain controversial significance level was much appreciated.

Mr. A. O. H. Roberts was a most patient and helpful consultant on statistical procedures and their psychometric implications and gave freely of his valuable time to exposition and critical discussion.

REFERENCES

BIESHEUVEL, S. (1942): No. 4 in a series, *Lectures to Senior Flying Instructors.* (Unpublished.)

KOERTH, W. (1922): A pursuit apparatus—eye-hand co-ordination. *Psychol. Monog.,* 31, 288-292.

WOODROW, A. P. (1956): Some observations on the two-hand co-ordination test as a predictor of flying ability. *J. Nat. Inst. Personnel Res.,* 7, 116-124.

Reprinted from the Journal of the National Institute for Personnel Research.

J. NAT. INST. PERSONNEL RES., 1961, 8, 199—208

HANDLEBARS: A SELF-PACED TEST OF TWO-HAND CO-ORDINATION, AND SOME RESULTS ON AIR-PILOT CANDIDATES

D. R. DE WET

(Received 31st December, 1960)

Among various sensory-motor tests applied to prospective air-crew in the South African Air Force during the last war, two on which considerable reliance came to be placed, were Steadiness and Two-hand co-ordination. This was founded both on the empirical validity of certain performance measures and also on the usefulness of these tests as situations favourable to the observation of wider subjective behaviour.

Steadiness involved direct tracing tasks, where a stylus-point was guided by one hand, firstly, down a straight tapering track, and then along a zig-zag. Two-hand co-ordination was of the well-known Moede type, in which the subject rotates the handles of two lead-screws located at right-angles to each other, to guide a metal platform carrying a card with a circular track, under a stationary pencil.

The "Handlebars" test to be described here is an apparatus designed to incorporate the following features: Firstly, two-hand co-ordination on a lever-system rather than that of rotating handles, as the former has more direct affinity with the main aircraft controls; and secondly, a "steadiness" element that is lacking in machines of the Moede type, where fine movements, voluntary or otherwise, do not register.

Description of Apparatus

(Plates 1, 2, 3 and 4)

The functioning principle of "Handlebars" is extremely simple, but careful fitting of the moving parts was necessary for smooth and precise operation.

A sloping wooden base-board supports a structure of metal plate attached by screws and spacing-pieces to a bakelite backing-sheet. The metal plate is cut out to form an open track, the contour of which is intended to provide within practicable limits, the maximum distance of continuous travel, and a wide variety of changes in direction.

Near the lower edge of the base-board, a sturdy vertical stub-axle is mounted; this forms the main pivot-point of two levers, each having a handle at one end and a bearing at the other coupled to a linking-bar. The other ends of the linking-bars are connected together to form a single bearing, from which a round rod protrudes downwards through the track in the metal plate, but has adequate clearance above the bakelite backing-sheet.

Fore-and-aft movement of the rod is obtained by drawing the handles nearer together or further apart; lateral movement, by pivoting them both simultaneously and drawing them together a suitable amount to compensate for the locus being an arc.

Duration of contact between the rod and the sides of the track is recorded on a "Standard" clock mounted on a separate control-box for the observer. One switch is in the main 220/230 volt A.C. circuit of the clock-motor, and the other, together with the rod and track, in the 6 volt A.C. circuit of the clock-clutch. A small transformer to supply the low voltage current, and a buzzer connected in parallel with the clock-clutch, are located inside the box.

Testing Procedure

The apparatus is placed on the table in such a way, that, with the metal rod at the starting-point (i.e., the upper section of the track to the right of the junction leading down to the middle or "star" section), the ends of the handles are in line with the table-edge where the subject will sit.

During 1954 and 1955 the following instructions were given to the subjects:

"This is a test of the smoothness and accuracy of your hand co-ordination. By using these two handles you have to steer the metal rod through the track, trying not to touch the sides. Starting from here, you go to the right round the outer track and then to the right round the star, finishing again at the starting-point. Whenever you allow the metal rod to touch the sides of the track, a buzzer will sound, which means that error is being recorded against you. Work carefully and as quickly as you can without touching the sides of the track. Now centralise the metal rod in the track so that it is not touching.

Are you ready? Begin!"

Subjects did two trials each, but they were not told of the second until they had completed the first. Measures were taken of both the "Total Time" and the "Error Time" for each trial.

During 1956 the number of trials was increased to five, and the new set of instructions read as follows:

"This is a test of your ability to steer a rod along a track. By using these two handles you have to steer the metal rod along the middle of this track. Starting from here, you go to the right round the outer track and then to the right round the star, finishing again at the starting point (Observer points round track). *Don't try to work fast!*

(1) This first round is a practice one. Work smoothly and carefully and try to keep this metal rod (Observer points to it) well in the middle of the track all the way along. Whenever you allow the rod to touch the sides of the track, a buzzer will sound, which means that errors are being recorded against you. Go round *very* cautiously and try your hardest to keep the rod in the middle of the track. You may begin!

(2) In this second round, speed is the important factor. Try to avoid touching the sides of the track, but drive as fast as you can. Ready? Begin!

(3) The third round is an easy one. I want you to do the test in the way that you feel suits you best. Some people prefer to work somewhat slowly and carefully, others feel more comfortable if they work a bit fast. Both can be "good types". Do this round in your own individual way. You may begin!

(4) You realise now that in this test the two important things are good speed and good steering so as not to touch the sides of the track. If you can do both well you will get a really fine score. This fourth round is very important. Work as fast as you can and try your best not to touch the sides. Ready? Begin!

(5) This is the fifth and last round and it counts the most points, so try to make it your best. Try to work even faster than you did in the last round, and make even fewer errors. Ready? Begin!"

(As before, "Total Time" and "Error Time" for each trial were recorded.)

Application and Results

The test was administered experimentally to pupil-pilot candidates for the South African Air Force, with the object of ascertaining its reliability and potential usefulness as a predictor of success or failure on courses of flying training.

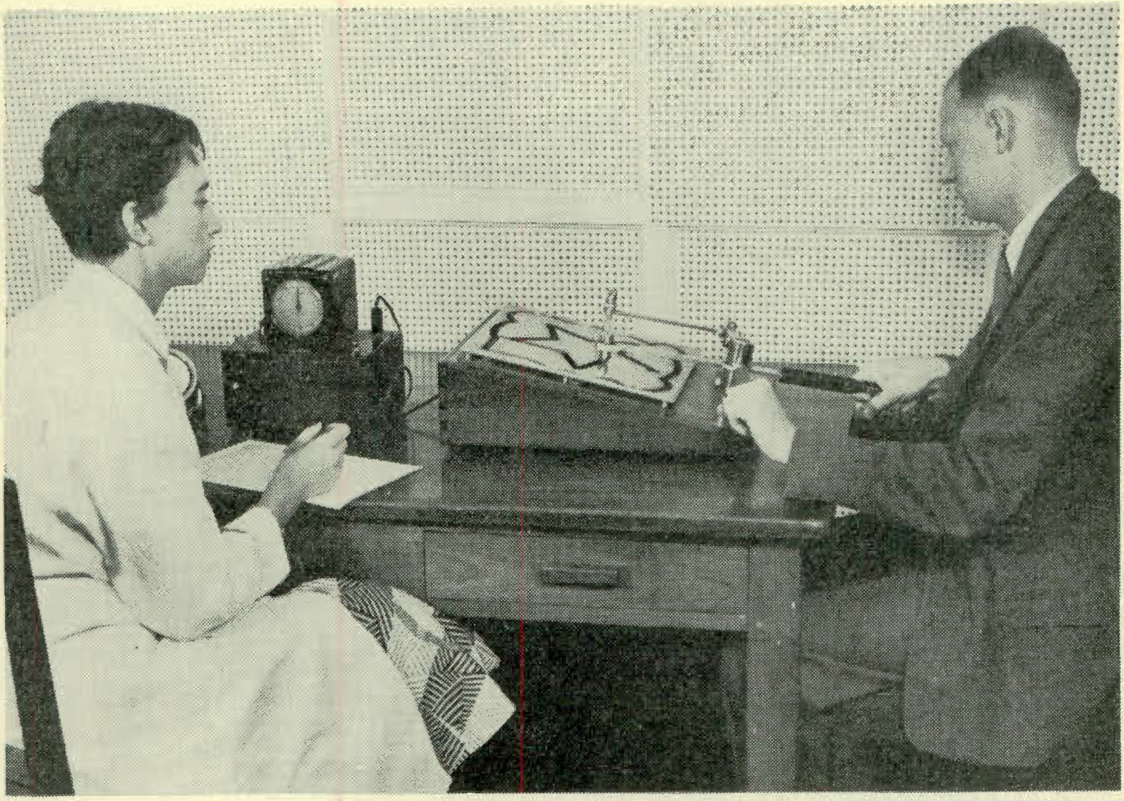
For validation purposes, successful pupils (Table I) and only those suspended for defective airmanship have been included; however, the means and inter-trial correlations for

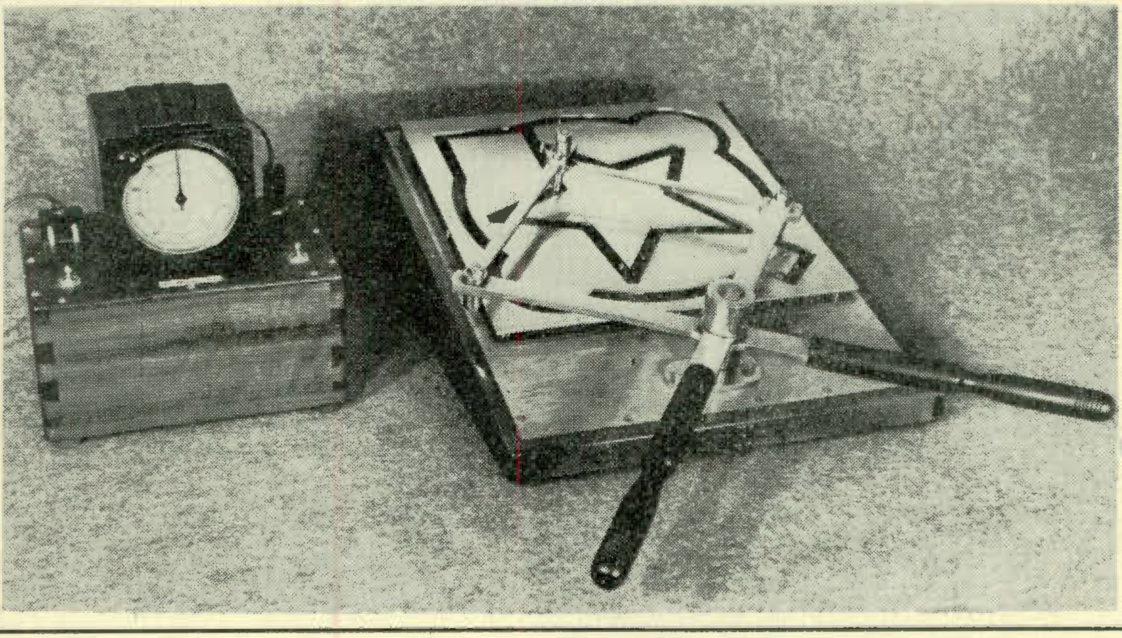
Table I

DERIVATION OF SAMPLES

	1954	1955	1956
Applied for flying course	93	169	163
Did Handlebars test	93	130	161
Accepted for training	42	51	47
Suspended from training (all reasons)	11	17	13
Suspended for poor flying ability	8	11	8
Did Handlebars test	8	11	8
Completed course successfully	31	34	33
Did Handlebars test	30	32	33

HANDLEBARS TEST. PLATE 1.





the 1956 group, include two additional pupils who were suspended for inability to master the theoretical training.

In Table II, the means for both "Correct Time" (i.e., Total Time minus Error Time) and "Error Time" reflect the influence of practice; there is the expected tendency to work faster and make less error on repetition of the task. In the 1956 group, the very conspicuous critical ratio of the difference between the "Correct Time" means of Trials 1 and 2, as compared with the ratios for Trials 1 and 2 of the other two groups, shows that in addition to practice there is the effect of the different instructions for these two trials,

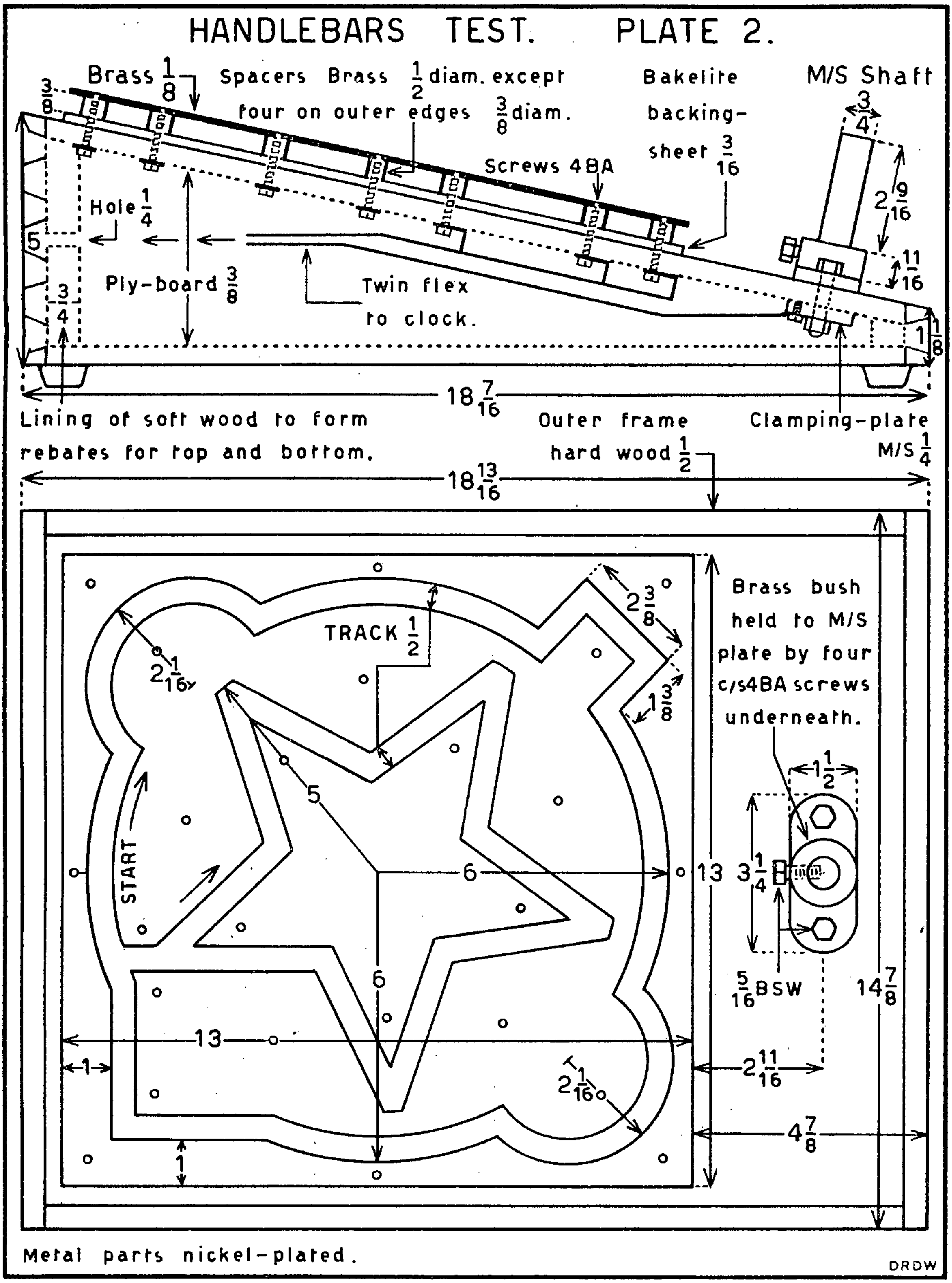
Table II

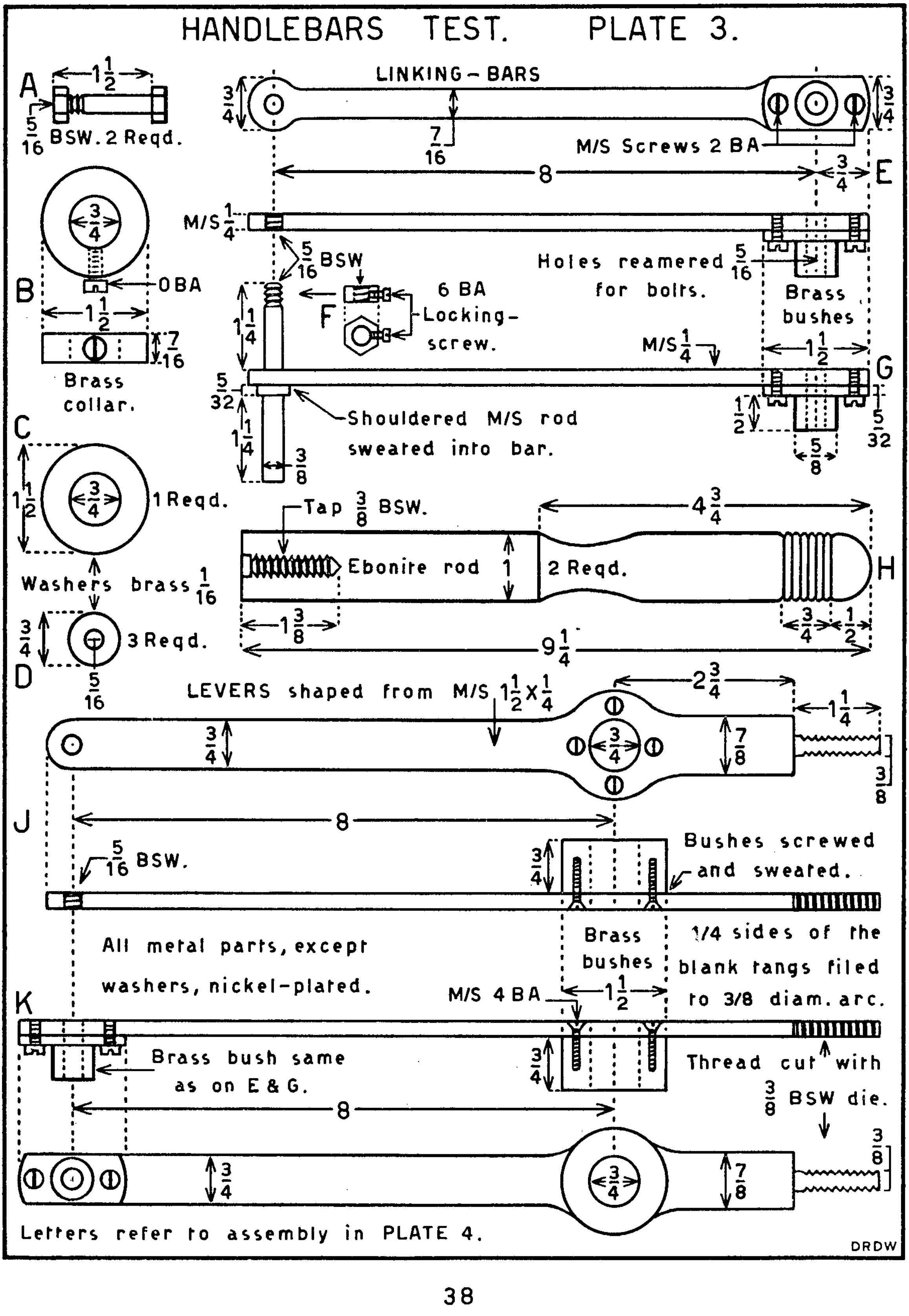
MEAN SCORES IN SECONDS

Pupil Group	No. of Cases	Trial No.	CORRECT TIME		ERROR TIME	
			Mean	S.D.	Mean	S.D.
1954	38	1	192.0	87.4	23.8	9.7
		2	157.1	78.0	18.9	5.9
1955	43	1	158.7	90.7	25.1	10.3
		2	128.6	68.2	21.1	7.6
1956	43	1	195.5	100.4	23.1	8.7
		2	102.6	78.8	18.9	5.7
		3	165.1	95.6	16.1	5.8
		4	94.7	64.2	15.0	4.3
		5	89.2	68.0	14.0	4.7

CRITICAL RATIOS OF THE DIFFERENCE BETWEEN THE MEANS OF TRIALS 1 and 2 FOR EACH OF THE THREE GROUPS

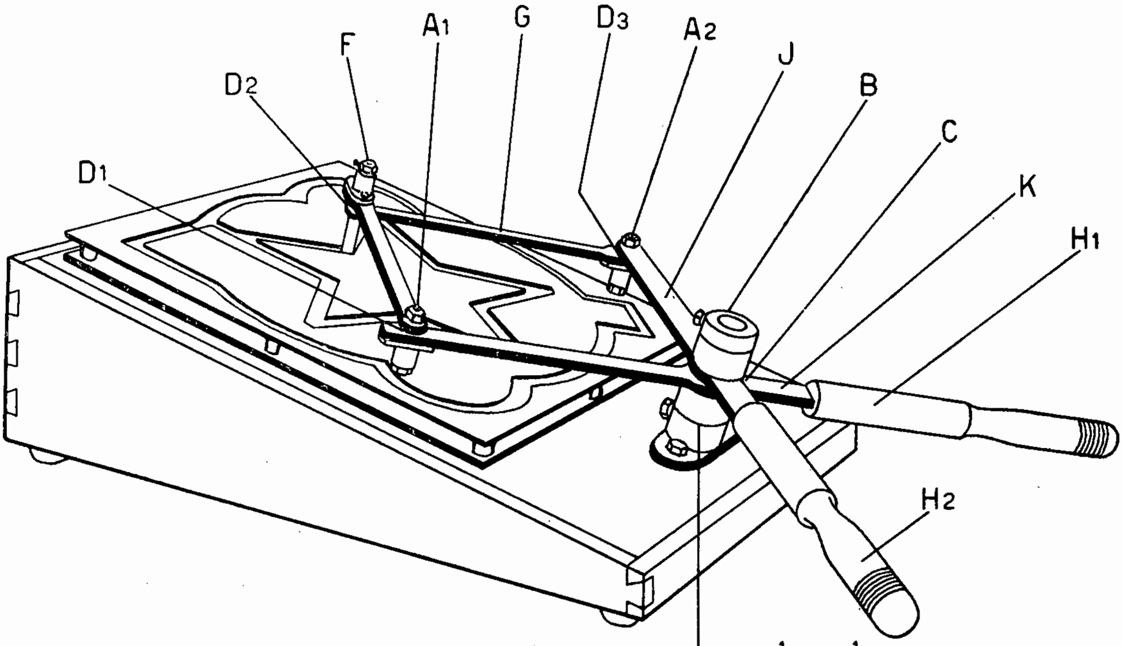
Group	Correct Time	Error Time
1954	4.46	4.32
1955	6.17	4.17
1956	11.99	7.92


HANDLEBARS TEST. PLATE 2.
Brass 1/8
Spacers Brass 1/2 diam. except four on outer edges 3/8 diam.
Bakelite backing-sheet 3/16
M/S Shaft
3/4
3/8
Screws 4BA
2 9/16
11/16
Hole 1/4
5
Ply-board 3/8
Twin flex to clock.
3/4
1 1/8
18 7/16
Lining of soft wood to form rebates for top and bottom.
Outer frame hard wood 1/2
Clamping-plate M/S 1/4
18 13/16
TRACK 1/2
2 3/8
2 1/16
1 3/8
Brass bush held to M/S plate by four c/s4BA screws underneath.
1 1/2
5
START
6
13
3 1/4
6
5/16 BSW
14 7/8
13
1
2 1/16
2 11/16
4 7/8
1
Metal parts nickel-plated.
DRDW


HANDLEBARS TEST. PLATE 3.
LINKING-BARS
M/S Screws 2 BA
A
5/16 BSW. 2 Reqd.
B
0 BA
Brass collar.
C
1 Reqd.
Washers brass 1/16
3 Reqd.
D
E
M/S 1/4
5/16 BSW
Holes reamered 5/16 for bolts.
Brass bushes
F
6 BA Locking-screw.
M/S 1/4
G
Shouldered M/S rod sweated into bar.
Tap 3/8 BSW.
Ebonite rod
2 Reqd.
H
LEVERS shaped from M/S 1 1/2 x 1/4
J
5/16 BSW.
Bushes screwed and sweated.
All metal parts, except washers, nickel-plated.
Brass bushes
1/4 sides of the blank tangs filed to 3/8 diam. arc.
M/S 4 BA
K
Brass bush same as on E & G.
Thread cut with 3/8 BSW die.
Letters refer to assembly in PLATE 4.
DRDW
38

HANDLEBARS TEST. PLATE 4.

(PLATE 3 REFERS)



Note: One purpose of the washers is to prevent identical metals from bearing against each other.

Polished M/S washer $1\frac{1}{2}$OD x $\frac{1}{64}$ between brass faces. (Not shown in PLATE 3)

DRDW

caution being stressed in the first and speed in the second.

It is of interest that in Trial 2 of the 1956 group where speed was stressed, significantly less error occurred than in the preceding trial with emphasis on caution. In the 1954 and 1955 groups, where the instructions stressed both speed and accuracy equally, the critical ratio of the difference between Trials 1 and 2 on Error Time was of the order of four; but in the 1956 group it is nearer eight and seems far too large to be accounted for by practice effect alone. The most likely explanation lies in some basic instrumental dependence that exists between Time and Error Time in self-paced tracing tasks, and has the effect of inducing a proportion of positive correlation between the two measures that tends to increase with speed of work. Opposed to this is the behavioural tendency for fast working to be associated with the production of more error, as shown in Table III. Here, for the 1956 group, the negative inter-correlations between "Correct Time" and "Error Time" are larger for Trials 1 and 3, done slowly, than for Trials 2, 4 and 5, done fast; this would support the hypothesis that the relative effect of instrumental dependence increases with speed of work.

In the 1954 and 1955 groups the smaller negative correlations between "Correct Time" and "Error Time" on the first trial, as compared with the second, suggest the influence of novelty in the test situation, where a certain amount of random fumbling by the tyro in his exploration for a successful technique reduces the consistence of the behavioural time-error relationship. The correlation for the first trial in the 1956 group has probably been depressed through the same influence. In the third trial, however, the subject has had the benefit of considerable practice, and a resultant stabilization of technique, conjoined with the important circumstance of unhurried work in this trial and hence less opposing effect through instrumental dependence, produces a

Table III

PRODUCT-MOMENT CORRELATIONS BETWEEN "ERROR TIME" AND "CORRECT TIME" (TOTAL TIME minus ERROR TIME) FOR EACH TRIAL.

Group	1954		1955		1956				
N	38		43		43				
Trials	1	2	1	2	1	2	3	4	5
r	— .34	— .49	— .33	— .47	— .40	— .33	— .58	— .31	— .27
Mean r = — .39								P < .001	

(Fast working on the test tends to be associated with more error)

Table IV

PRODUCT-MOMENT INTERCORRELATIONS BETWEEN TRIALS ON THE MEASURES "CORRECT TIME" AND "ERROR TIME"

GROUP	TRIALS	CORRECT	ERROR	N
1954	1 & 2	.84	.71	38
1955	1 & 2	.96	.80	43
1956	1 & 2	.87	.64	43
	1 & 3	.80	.68	
	1 & 4	.73	.65	
	1 & 5	.76	.65	
	2 & 3	.77	.59	
	2 & 4	.84	.84	
	2 & 5	.77	.89	
	3 & 4	.80	.65	
	3 & 5	.75	.65	
	4 & 5	.95	.86	
	Mean r	.84	.73	124
P of the difference = .020				

comparatively large correlation. A further investigation of the interesting relationship between "Correct Time" and "Error Time" (or "Total Time" and "Error Time") on self-paced tasks, will be worthwhile.

In Table IV, allowance must be made for the fact that most of the intercorrelations between trials in the 1956 group have probably been depressed through different instructions, and overall mean coefficients are therefore an under-estimate of test reliability. A better indication of reliability is given in Table V, which is restricted to those trials where instructions were similar, namely, equal emphasis on speed and accuracy. Both tables show

Table V

PRODUCT-MOMENT INTERCORRELATIONS BETWEEN THOSE TRIALS HAVING SIMILAR INSTRUCTIONS

GROUP	TRIALS	CORRECT	ERROR	N
1954	1 & 2	.84	.71	38
1955	1 & 2	.96	.80	43
1956	4 & 5	.95	.86	43
	Mean r	.93	.80	124
P of the difference < .001				

the measure "Correct Time" to have more reliability than "Error Time", the difference being more significant in Table V, where, in addition to uniformity of procedure, the close adjacency between these particular pairs of trials makes for larger correlations.

Table VI

BISERIAL CORRELATIONS BETWEEN BOTH "CORRECT TIME" AND "ERROR TIME" AND SUCCESS/SUSPENSION IN FLYING TRAINING

GROUP	1954		1955		1956				
N	38		43		41				
TRIALS	1	2	1	2	1	2	3	4	5
CORRECT T r bis	.21	.18	.06	.16	.32	.20	.04	.26	.27
ERROR T r bis	— .03	— .01	— .14	.15	— .35	— .29	— .27	— .25	— .11

P (one-tailed) of .32 = .039

P (one-tailed) of .35 = .030

Table VII

PRODUCT-MOMENT CORRELATIONS BETWEEN THE HANDLEBARS TEST AND THREE OTHER SENSORY-MOTOR TESTS IN THE 1956 GROUP

HANDLEBARS TOTAL TIME:		r	P
Two-hand Co-ordination (Moede type).	Time	.152	.055
Hand-foot Reaction (Speed trial).	Time	.035	.660
Steadiness (Manual).	Total Time	.532	< .001
HANDLEBARS ERROR TIME:		r	P
Two-hand Co-ordination (Moede type).	Errors	.219	.006
Hand-foot Reaction (Speed trial).	Errors	— .106	.180
Steadiness (Manual).	Error Time	.404	< .001

Number of cases = 161. (P is two-tailed).

In Table VI, positive signs on "Correct Time" imply that success in flying training tends to be associated with slow working on the test, and negative signs on "Error Time" that success tends to be associated with less error on the test. For the most part correlations between test performance and the flying-training criterion are not significant, an exception, however, being the first trial in the 1956 group, where correlations on both measures are significant at better than the 4 per cent level. While results on the Two-hand co-ordination (Moede type, 1956), Floating Effect (1958), and Steadiness (1959) tests suggest that the first trial in certain sensory-motor activities seems to have some special predictive value because of its novelty, the significance of the first trial on the Handlebars test for the 1956 group could also be related to the emphasis on slow and careful work. This possibility is strengthened by the small but fairly consistent tendency for slow work to be associated with success, even in trials where the instructions stress fast work.

The correlations between performance on Handlebars and three other sensory-motor tests (Table VII) are largest and very significant for Time and Error Time on the Steadiness test (1959), and a good deal smaller, but still acceptably significant, for "Time" and "Errors" on the Two-hand co-ordination test (Moede type, 1956). No relationship is shown between Handlebars and the Hand-foot Reaction test (1958).

Summary and Conclusions

(1) Details are given of the design and experimental application to air-pilot candidates of a two-hand co-ordination apparatus functioning on a lever-system and incorporating an element of "steadiness".

(2) As a measure of test performance "Correct Time" (i.e., Total Time minus Error Time) has more reliability than "Error Time".

(3) Fast working on the test is associated with more error, to the extent of a mean correlation of — .39 (P = .00006) between "Correct Time" and "Error Time", which has probably been depressed through the opposite influence of some positive instrumental dependence between these variates. Such an effect can be sensed on other self-paced tracing tasks, and in the case of Steadiness (1959) the occurrence of several positive correlations between "Correct Time" and "Error Time" suggests that is was particularly significant.

(4) "Handlebars" bears a closer relationship to the Steadiness test (1959) than to a Two-hand co-ordination test of the Moede type (1956).

(5) There is a small but fairly consistent tendency for slow working and less error on the test to be associated with success in flying training; but in this respect, the only trial having individual significance is Trial I in the 1956 group, where the instructions emphasised slow and careful work. Here the biserial correlations are: "Correct Time", .32 (P = .039), and "Error Time", .35 (P = .030).

(6) The combination in a single apparatus of certain elements of a Steadiness test, and a two-hand control-system on a lever-principle that seemed a closer approach to the feel and movement of actual aircraft controls than that given by the rotating handles on machines of the Moede type, has not resulted in a significantly better instrument for the prediction of flying ability. It could possibly be improved by instituting a stress trial in which, by means of a pair of light electrodes strapped to the subject's arm, he would either work under the threat of shock at the end of the trial (as in the Steadiness test), or actually receive shock whenever he made mistakes.

(7) Although it appears to have small validity for air-pilots the Handlebars test might be usefully applicable in other fields, such as the skilled and semi-skilled artisan trades, or various industrial occupations requiring proficiency in the use of machine tools.

Acknowledgments

Mr. J. B. Kirstein took the photographs. The test was administered by Mrs. A. P. van der Reis, Mrs. J. du Plessis and other members of the N.I.P.R.

REFERENCES

DE WET, D. R. (1958): Co-ordination and floating effect. *J. Nat. Inst. Personnel Res.*, 8, 28-38.

DE WET, D. R. (1958): A portable hand-foot reaction test. *J. Nat. Inst. Personnel Res.*, 8, 106-116.

DE WET, D. R. (1959): An improved steadiness apparatus and its validity for air-pilot selection. *J. Nat. Inst. Personnel Res.*, 8, 122-136.

WOODROW, A. P. (1956): Some observations on the two-hand co-ordination test as a predictor of flying ability. *J. Nat. Inst. Personnel Res.*, 7, 116-124.

Reprinted from "Psychologia Africana," Volume 9, 1962.

86 PSYCHOLOGIA AFRICANA, 1962, 9, 86–99

A VARIABLE CO-ORDINATION TEST AND ITS POTENTIALITY AS A GAUGE OF APTITUDE FOR AIRMANSHIP

D. R. DE WET

Received 24th April, 1961

No test has yet been evolved that can be described as a universal tool, equally suitable for application in every field in which muscular co-ordination plays a substantial part; like the mythical philosopher's stone, such a device will probably remain an unrealized ideal. While there is evidence of group factors overlapping in certain sensory-motor skills, in others specificity appears predominant, and the demand for specific selection tests continues.

The closer the behavioural correspondence or " feel " between test and actual occupation, the more useful the former is likely to be as a predictor of occupational success, but the genuineness of this similarity is not necessarily guaranted by exact reproduction in the test of certain obvious material characteristics of the original. Such imitation may well produce a spurious laboratory model that bears a superficial resemblance to actuality, but lacks the psychological essentials. The latter are the all-important ingredients, and provided that *their* presence is assured in the test then the addition of other similarities, which would be worthless on their own, may contribute something of value by completing the total verisimilitude.

As there would be no advantage whatever in a laboratory facsimile as complicated as the real occupation, the striving is towards simplification, by concentrating attention on those aspects which seem to be most important. With regard to both basic interest, and considerations of practicability and economy in applied fields, there is a strong incentive to investigate and exploit to the full whatever potentialities might lie in group factors of ability.

For personnel selection, the criterion of parsimony that governs " intellectual " and similar tests also applies in designing tests of sensory-motor capacities. Notwithstanding the great preponderance of specific matter in this last category, research here is likewise directed towards achieving an optimum situation where one condensed test adequately covers the common requirements for several occupations, and has to be supplemented by specific tests as few in number and elementary in structure as possible.

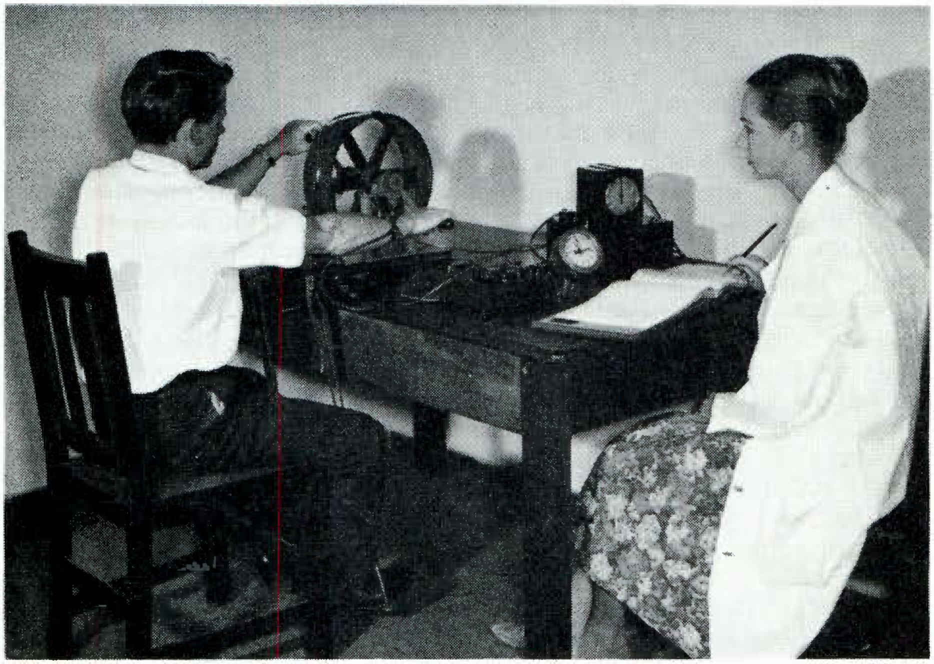
The Variable co-ordination test that will be dealt with here is an attempt to produce a piece of equipment that lends itself to both basic research on co-ordination and occupational selection. In the mechanical design, careful consideration has been given to such points as reasonable portability, the minimum number of moving parts, the use of ordinary commercial items wherever practicable, and overall sturdiness to prevent loss of standardization through wear or through possible damage to components as a result of rough treatment by the subjects.

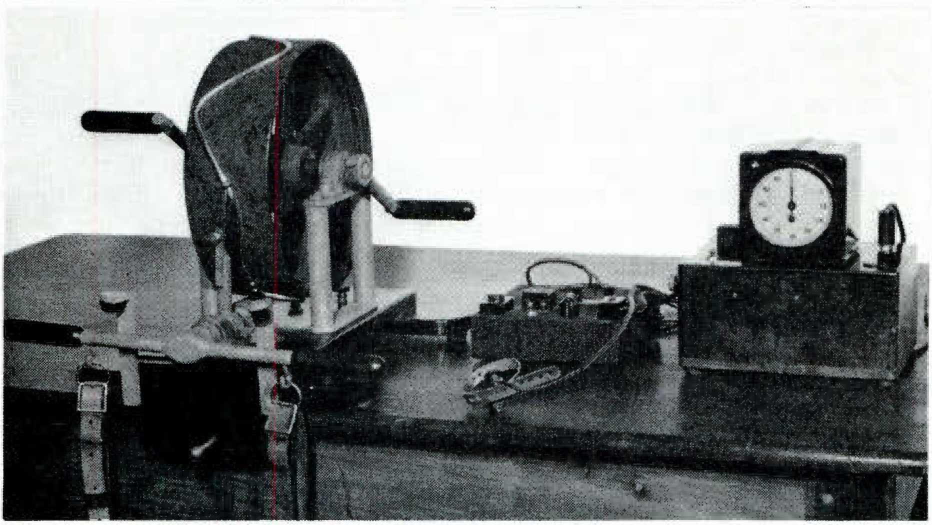
Description of apparatus (Plates 1, 2, 3, 4 and 5)

Fundamentally the apparatus consists of two moving parts.

(i) A wheel, with a raised track of irregular contour fixed around its periphery, is mounted between two bearings on a shaft having a crank handle at either end.

VARIABLE CO-ORDINATION TEST. PLATE 1.





(ii) Another pair of bearings, situated lower and at right angles to the wheel bearings, supports the shaft of a horizontal crossbar and a vertical arm terminating in a small knob on a spring-wire stem. The crossbar is fitted with handles, and also two hooks from which, by means of leather straps and rings, a pair of ordinary equestrian stirrup irons may be suspended. To permit the use of one stirrup only, or the introduction of loading to alter the trim of steering systems employing both stirrups or none, a small bollard on the front of the chassis serves for mooring one end of a spiral spring linked at its other end to the appropriate hook (Plate 3).

There are two separate control boxes for the observer (Plate 1). The one carries a "Standard" clock, transformer, triple-pole relay, buzzer, and switch; the other is a shocking-unit identical in principle with that built into the Steadiness test (2). As one requirement in the subject's task is that the small knob should be kept in contact with the track, these two sections together constitute a switch in series with the relay solenoid; while the latter is kept energized by contact, the buzzer remains silent, and no recording takes place on the clock (Plate 5). However, when the knob goes off the track, the relay armature is released and both clock and buzzer come into operation.

The middle pair of relay contacts is connected to the shocking-unit as shown in Plate 5, so that with a pair of electrodes strapped to the subject's arm, in addition to the sounding of the buzzer and the recording of error time on the clock, shock is automatically administered while the knob is off the track. Although the recording and shocking systems are treated together in Plate 5, in practice (Plate 1), the latter has been constructed as a separate unit to make it conveniently adaptable for use with other tests. A push button "C" located on the shocking-unit and shown wired in parallel with the lead from the relay contacts in the box supporting the clock, is intended mainly for manual administration of shock by the observer in other test situations where it may be required. In the present situation it provides a quick means for the observer to do occasional checks on the intensity of any shock setting by holding the electrodes against his own arm and pushing the button.

Testing procedure

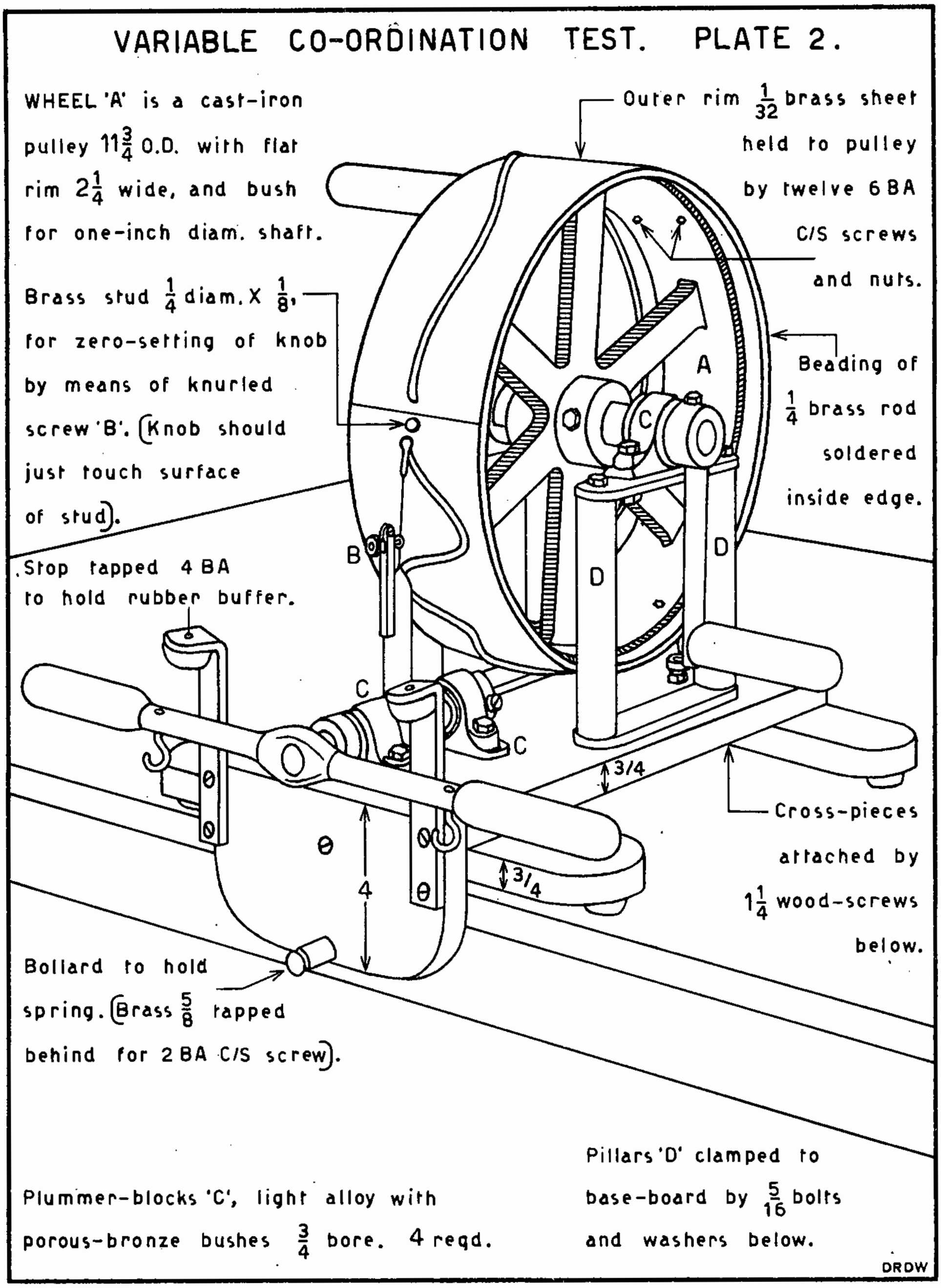
The task in this test is of the self-paced type, the subject being required both to rotate the wheel and steer the knob on the track. Its "variability" refers to the several different combinations that can be imposed for the accomplishment of this task. For all of them, the general positioning of the apparatus on a table as shown in Plate 1 is the same, the small apron section at the front of the chassis always being up against the edge of the table.

The following co-ordinative combinations can be obtained.

Section A

Balanced steering mechanism

	1	2	3	4	5	6	7	8	9	10	11
Rotation of wheel ..	RH	LH	LH	RH	LH	RH	LH	RH	BH	BH	BH
Steering of knob ..	LH	RH	RF	LF	LF	RF	BF	BF	RF	LF	BF


VARIABLE CO-ORDINATION TEST. PLATE 2.
WHEEL 'A' is a cast-iron pulley 11 3/4 O.D. with flat rim 2 1/4 wide, and bush for one-inch diam. shaft.
Outer rim 1/32 brass sheet held to pulley by twelve 6 BA C/S screws and nuts.
Brass stud 1/4 diam. X 1/8, for zero-setting of knob by means of knurled screw 'B'. (Knob should just touch surface of stud).
Beading of 1/4 brass rod soldered inside edge.
A
B
C
D
Stop tapped 4 BA to hold rubber buffer.
3/4
4
Cross-pieces attached by 1 1/4 wood-screws below.
Bollard to hold spring. (Brass 5/8 tapped behind for 2 BA C/S screw).
Plummer-blocks 'C', light alloy with porous-bronze bushes 3/4 bore. 4 reqd.
Pillars 'D' clamped to base-board by 5/16 bolts and washers below.
DRDW

Section B

Unbalanced steering mechanism

On certain of the above, additional variations in the distribution of muscular effort are obtainable by using the spring to load the steering system in one direction or the other.

	12	13	14	15	16	17	18	19	20	21
Rotation of wheel	RH	RH	LH	LH	LH	LH	RH	RH	BH	BH
Steering of knob	LH	LH	RH	RH	BF	BF	BF	BF	BF	BF
Spring between bollard and hook	L	R	L	R	L	R	L	R	L	R

R = Right; L = Left; H = Hand(s); F = Foot (Feet); B = Both.

In the present investigation, only combination No. 11 involving both hands and both feet (balanced steering) was used, but with a few suitable changes the procedure and instructions will also serve for any of the others:—

The subject is seated at such a distance from the apparatus that, with his left arm fully extended he can just grasp the left handle of the wheel at its furthest position. His feet are then settled in the stirrups slightly forward of the insteps, and the wheel is turned to the "start" position with the stud (Plate 2) opposite the knob. He is then instructed as follows:

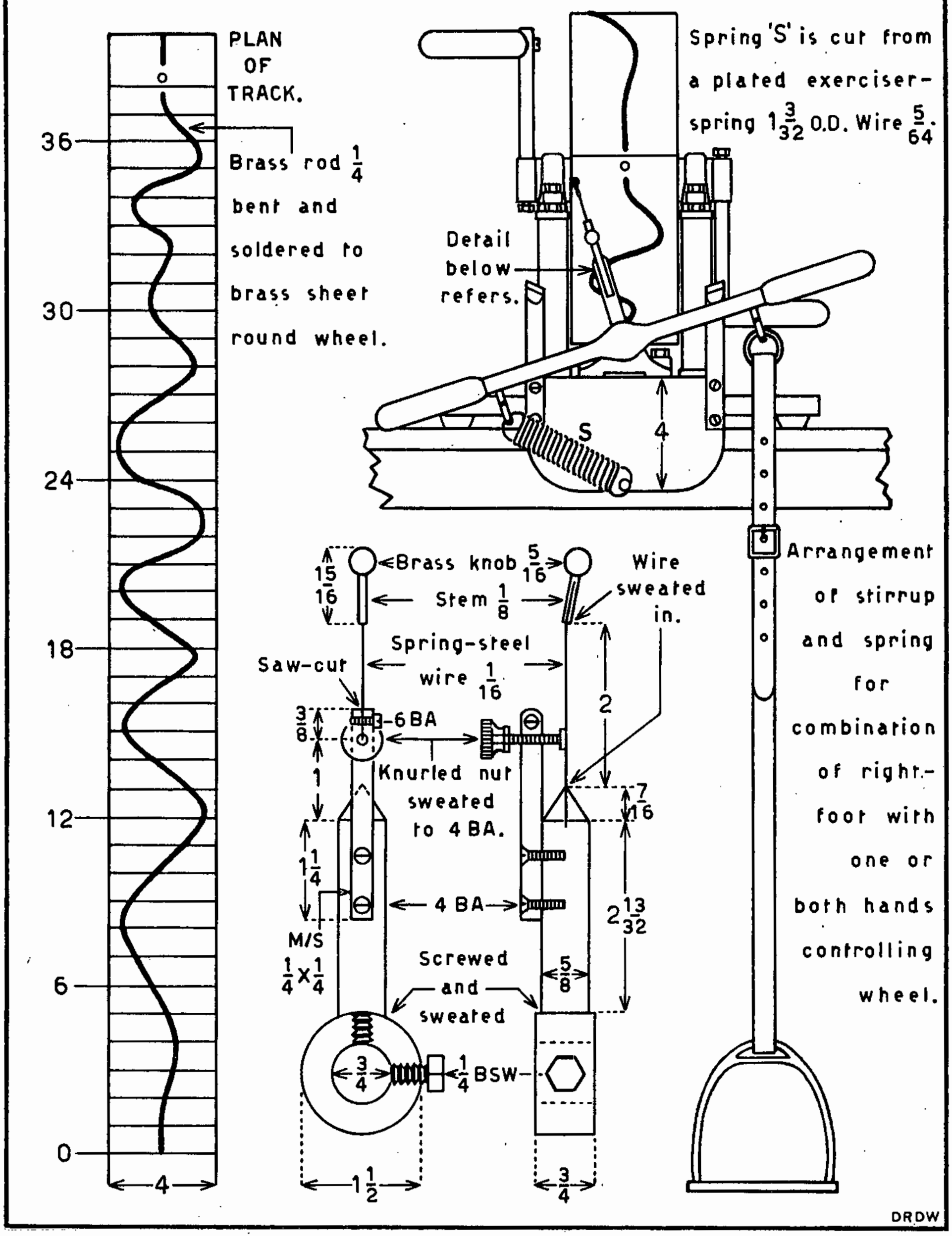
"This is a test of your co-ordination and smoothness of movement. By using the stirrups, this knob may be steered to the left or right. With your hands on these handles, your task is to turn the wheel carefully towards you like this, and at the stame time use your feet to keep the knob in contact with the track. Whenever you allow the knob to get off the track, you will hear a click and a buzzing noise. This means that error is being recorded against you, so get back on to the track as soon as you can. You must concentrate on trying to work both fast and carefully. Remember, your accuracy in keeping to the track and the time you take to go round it are both equally important. Get your hands on the handles and have the wheel at the starting position there. Don't start until I tell you to. Ready? Start!" [A stop-clock and the main switch "A" (Plate 5) are turned on.]

When the subject reaches the end of the track, stop-clock and switch "A" are turned off and he is told to "Rest." The total time he has taken to complete the circuit, and also his "time off the track" as registered on the "Standard" clock, are recorded.

After a rest pause of 30 seconds, the usual instruction is: "You start another trial now; try to improve. Ready? Start!" However, if the first trial was performed in a very hurried fashion with obvious neglect of accuracy, the subject is now told: "Try to make fewer mistakes." On the other hand, if he has taken excessively long to complete the first trial, he is told: "Try to work faster."

The second circuit is also followed by a rest pause of 30 seconds, "Total Time" and "Error Time" being recorded as before. The shock-electrodes are then fitted to the outer side of the subject's arm about three inches behind the wrist-joint. They should be in close contact with the skin, but not too tight. "This is the third trial; whenever you go off the track now, in addition to hearing the noise, you will feel a shock. Just keep on trying to improve. Ready? Start!" (Switch "B" must not be turned on until the knob is in contact with the beginning of the track. The trial is followed by a rest pause, and the scores are recorded. Before the last trial is started, the potentiometer on the shocking-unit is set to a

VARIABLE CO-ORDINATION TEST. PLATE 3.



point that raises the previous intensity of the shock by about a half.) "This is the last trial, try to make it your best. When you get off the track now, the shock will be stronger. Ready? Start!" (The final set of scores is recorded in the same way.)

Application and results

A group of air-pilot candidates for the South African Air Force was given the test experimentally on the combination involving both hands and both feet (balanced steering). The purpose of this initial study was to investigate the test's reliability, and its value as an instrument for assisting in the future selection of candidates most likely to succeed in airmanship.

The sample dealt with here is made up of the two smaller groups (a) and (b) in Table 1, and confined to candidates who were accepted for training. Suitability for admittance to the course was determined through findings on other testing and interview procedures, and not influenced by performance on the present test. Six candidates who were suspended for reasons other than that of poor flying ability, have been excluded.

Table 1

Derivation of sample

Year	1959
{ Applied for flying course	152
{ Did Variable co-ordination test	149
Accepted for training	55
Suspended from training (all reasons)	22
{ Suspended for poor flying ability	16
{ Did Variable co-ordination test	16(b)
{ Completed course successfully	33
{ Did Variable co-ordination test	31(a)

In Table 2, the means for "Error Time" show a progressive diminution through the four successive trials, and clearly indicate the effect of practice. For "Correct Time" (i.e. "Total Time" minus "Error Time") the position is somewhat different. Here there is improvement in speed of work between Trials 1 and 2, and then in Trial 3, notwithstanding the effect of previous practice, a very significant regression to slow working brought about through the fear of punishment by shock whenever error is made.

It was noticeable during the administration of the test that some of this fear seemed to be anticipatory, and subjects usually started off very slowly in Trial 3 and continued thus until they received the initial shocks. Henceforth, the realization that it was not as painful as they had expected lessened apprehension somewhat, and they risked stepping up the pace. As a result of this reassurance, and also the considerable degree of physical adaptation to actual shock that took place through repetitive experience, the effect of practice is allowed to emerge again in Trial 4, where there is a tendency for speed to show an improvement over that in Trial 3 despite an increase of shock intensity. But in Trial 4 the fear of punishment as an incentive to slow and careful work, although in a measure reduced by familiarity and adaptation during the course of Trial 3, still persists

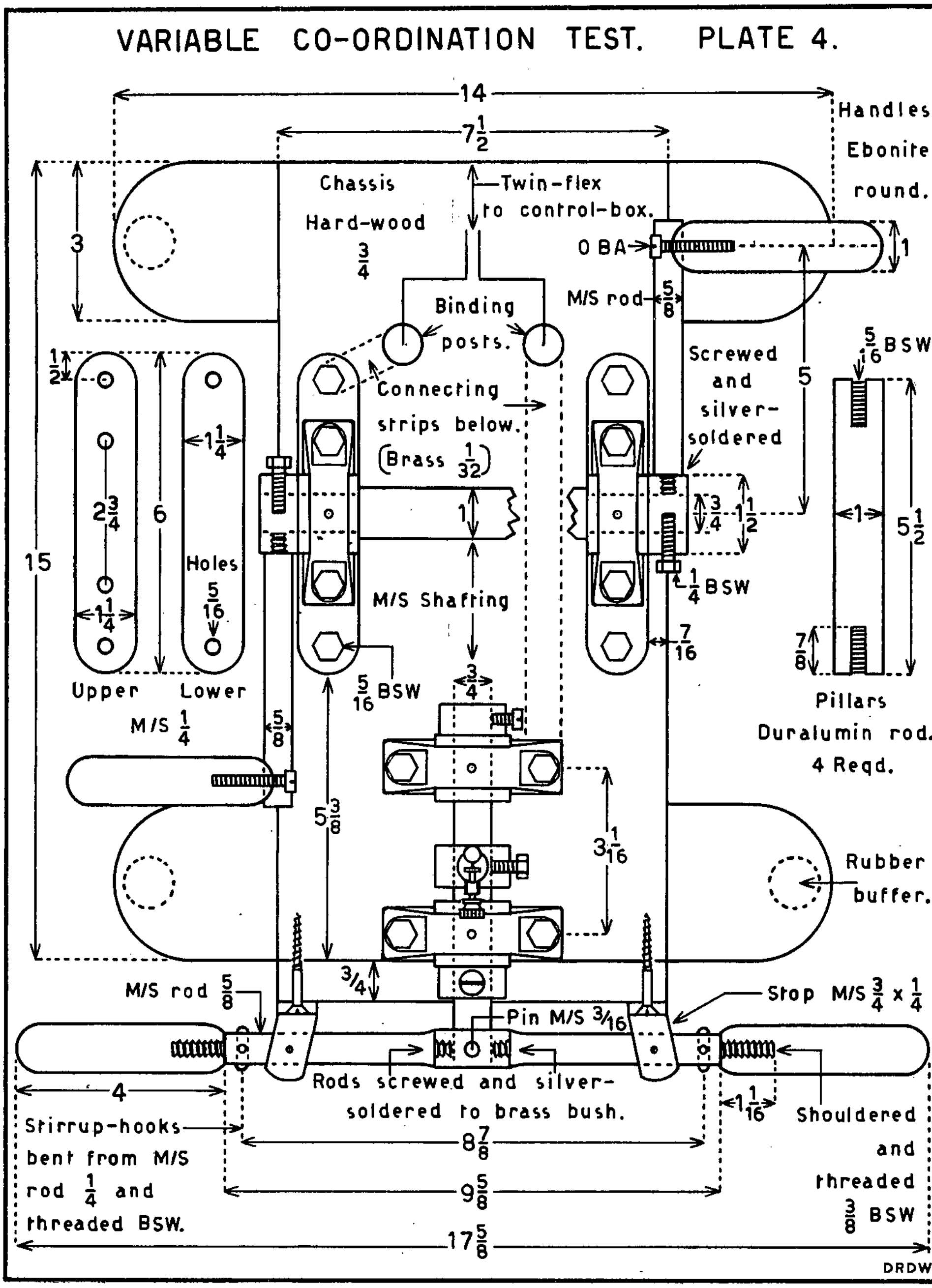

VARIABLE CO-ORDINATION TEST. PLATE 4.
14
7½
Handles
Ebonite
round.
Chassis
Hard-wood
¾
Twin-flex
to control-box.
0 BA
1
M/S rod 5/8
Binding
posts.
Connecting
strips below.
(Brass 1/32)
Screwed
and
silver-
soldered
5
5/16 BSW
3
½
2¾
6
1¼
Holes
5/16
15
1¼
1
¾
1½
1
5½
M/S Shafting
¼ BSW
7/16
7/8
Upper
Lower
M/S ¼
5/8
5/16 BSW
¾
Pillars
Duralumin rod.
4 Reqd.
5⅜
3 1/16
Rubber
buffer.
M/S rod 5/8
¾
Stop M/S ¾ x ¼
Pin M/S 3/16
Rods screwed and silver-
soldered to brass bush.
4
1 1/16
Stirrup-hooks
bent from M/S
rod ¼ and
threaded BSW.
8⅞
9⅝
17⅝
Shouldered
and
threaded
⅜ BSW
DRDW

Table 2

Mean scores in seconds

Trial No.	Correct Time		Error Time	
	Mean	S.D.	Mean	S.D.
1	119·6	54·6	24·1	10·7
2	109·9	48·7	16·6	10·3
3	163·2	72·3	12·0	7·5
4	141·6	65·4	10·1	7·5

Critical Ratios of the differences between the above means

Trials	Correct Time	Error Time
1 & 2	2·21	4·97
1 & 3	6·79	8·18
1 & 4	3·79	9·52
2 & 3	7·87	3·38
2 & 4	6·12	5·56
3 & 4	4·26	2·35

Number of cases = 47

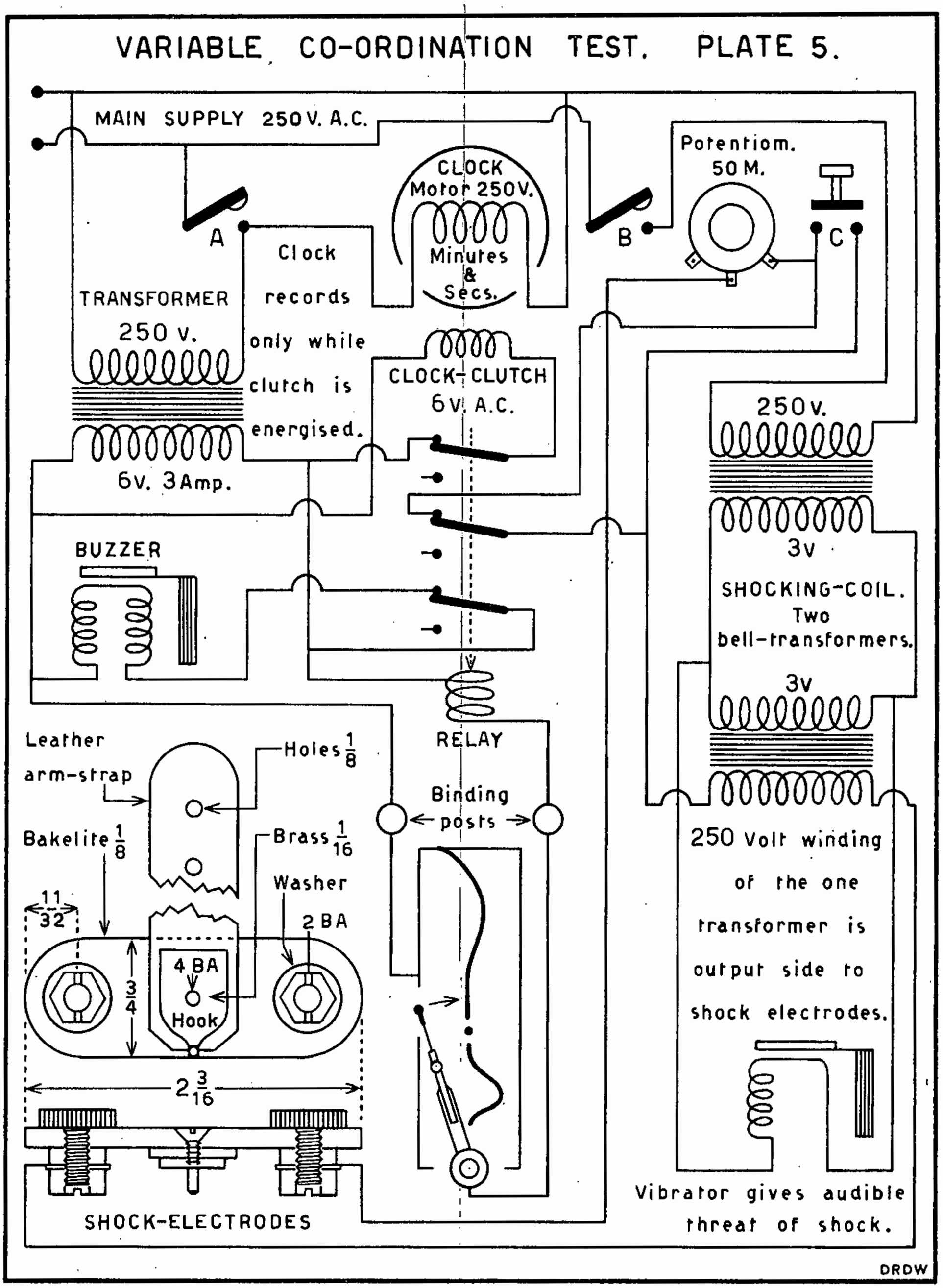
strongly and is reflected in the significant difference between the mean for Trial 4 and the means for the unstressed first and second trials, particularly the latter.

The correlations between " Correct Time " and " Error Time " for each trial, given in Table 3, show a general tendency similar to that observed in relation to performance on the Steadiness (2) and Handlebars (3) tests. Negative correlations are larger for the trials done slowly (in this case, 3 and 4) than for those done fast (1 and 2). This would also be compatible with the hypothesis that the factor of basic instrumental dependence that exists between " Time " and " Error Time " in self-paced sensory-motor tasks, becomes more prominent as speed of work increases. It is notable that Trial 4 has a more substantial negative correlation than Trial 3, although time for the latter was longer (Table 2). It seems likely that at this advanced stage of the test the effect of instrumental dependence, with its bias towards the production of positivity in the correlation between " Time " and

Table 3

Product-moment correlations between " Error Time " and " Correct Time " (" Total Time " minus " Error Time ") for each trial

N	47			
Trials	1	2	3	4
r	–·21	–·22	–·36	–·54
P	·16	·14	·02	·0004


VARIABLE CO-ORDINATION TEST. PLATE 5.
MAIN SUPPLY 250 V. A.C.
CLOCK Motor 250V.
Minutes & Secs.
Potentiom. 50 M.
A
B
C
Clock records only while clutch is energised.
TRANSFORMER 250 V.
6V. 3 Amp.
CLOCK-CLUTCH 6 V. A.C.
250 V.
3v
SHOCKING-COIL. Two bell-transformers.
3v
BUZZER
RELAY
Leather arm-strap
Holes 1/8
Bakelite 1/8
Brass 1/16
Washer
2 BA
4 BA
Hook
11/32
3/4
2 3/16
Binding posts
250 Volt winding of the one transformer is output side to shock electrodes.
Vibrator gives audible threat of shock.
SHOCK-ELECTRODES
DRDW

" Error Time," is overrun by other factors such as improvement in motor skill through practice and an increased familiarity with the " punishment " situation. These together would tend to make the subject's technique more stable, and thereby boost the negativity of the correlation.

Reliability of the test, as indicated by the intercorrelations between trials in Table 4, appears significantly higher for " Correct Time " than for " Error Time," which is in agreement with results obtained on the Steadiness and Handlebars tests. A probable explanation lies in the fact that in all three tests " Error Time " usually constitutes a much smaller proportion of the total than " Correct Time," and is therefore more markedly influenced by chance fluctuations both in the subject's response pattern, and in the mechanism of the apparatus where minor irregularities like specks of dust or tarnish between contacting surfaces can alter the sensitivity of the recording.

Table 4

Product-moment inter-correlations between trials on the measures " Correct Time " and " Error Time "

Trials	Correct Time	Error Time
1 & 2	·84	·53
1 & 3	·80	·44
1 & 4	·80	·45
2 & 3	·78	·50
2 & 4	·85	·64
3 & 4	·88	·73
Mean r's	·84	·55
P of the difference $<$ ·005		
N=47		

The relationship between scores on the test and proficiency in flying training is shown in Table 5. Instead of the usual method of biserial correlation on successes and suspensions, the more refined rank-difference procedure was applied to this particular group. Successful pupils were ranked in accordance with their final positions on the flying course. These were followed by the failures, ranked from the pupil suspended last to the pupil suspended first.

On the present test, contrary to the tendency observed in the Steadiness and Handlebars tests, *fast* working is associated with success in flying training. Inclining towards agreement with results on Handlebars there is indication that, to a small extent, less error favours the chance of success; whereas error on the Steadiness test showed no correlation with the flying criterion. This is consistent with the idea that the kind of co-ordinative activity in the Handlebars test and the full hand-foot combination (balanced steering) on the Variable co-ordination test, is less remote from aircraft-pilotage than that incorporated in the Steadiness test, where the small relationship with pilot success probably concerned rate of work as a temperamental rather than a skill factor.

Table 5

Rank-difference correlations between final placement of pupils on the flying course, and scores on both " Correct Time " and " Error Time "

Trial No.	Correct Time		Error Time	
	ρ	P	ρ	P
1	·34	·03	−·23	·13
2	·28	·08	·09	·57
3	·42	·01	−·25	·11
4	·48	·002	−·14	·35
N=47				

(Positive signs on " Correct Time " and negative signs on " Error Time " imply, respectively, that success in flying is associated with fast working and less error on the test).

" Correct Time " in Table 5 appears to relate more closely to position attained on the flying course than " Error Time." Of the four trials the two last involving punishment of error by electric shock are the most significant, particularly the fourth with more severe shock.

Below are two supplementary rank-difference correlations, obtained on the measure " Correct Time," between final positions of pupils on the flying course and the following derived scores.

(i) The difference between the average for the two unstressed trials (1 and 2) and the average for the two stress trials (3 and 4):

$\varrho = 0{\cdot}42 \qquad P = 0{\cdot}01.$

(ii) The score for Trial 4 plus the above difference:

$\varrho = 0{\cdot}49 \qquad P = 0{\cdot}001.$

These correlations imply that the smaller the increase of time under stress conditions, the better the chance of success, and it would appear that in addition to the motor aspects of the test, there is a temperamental element in the form of fear of punishment that also relates to proficiency in flying. This is supported by a finding on the Steadiness test that the trial carrying the threat of shock had more significance in relation to flying success than the unstressed trials.

Aircraft-pilotage continues to maintain its tradition of being a comparatively hazardous occupation; and to a degree that varies with individuals, it fosters the dread that mistakes can lead to either physical disaster or, at the very least, injury to the pupil's ego through adverse criticism by the instructor. A well balanced sensibility to danger can be regarded as an excellent and very necessary attribute for success in airmanship, but obsessive fears of pain or disgrace would certainly incommode the pupil and reduce his chance of succeeding.

While subjects were doing the test, the observer made notes of their general comportment and any particular reactions that might have significance. Points of interest that emerged from a later examination of these " reports " with reference to the pupils who succeeded in learning to fly and to those who failed, are given in Table 6.

Table 6

Some points of behaviour observed during the test, and the percentages of the 31 successful pupils (S), and the 16 failures (F), who manifested them

	Percentage Approx.		P of Diff.
	S	F	
(1) Very willing and co-operative	32	25	·61
(2) Fidgety and restless	23	19	·75
(3) Movements jerky, and/or a tendency to lead with one set of controls and follow with the other	26	44	·22
(4) Extremely talkative and socially at ease	23	6	·08
(5) Very active and hurried throughout	16	6	·25
(6) Overtly responsive to potential and/or actual shock	10	44	·01
(7) At some stage or other pressed so hard on the stirrups that the back of the machine tilted up off the table	23	6	·08

Fear of shock, positively associated with failure, stands out as the most significant (P = 0·01). Talkativeness and sociability (P = 0·08), and the tendency to press very hard on the stirrups (P = 0·08), are both positively related to success. Evidently this last item must not be identified with the kind of "tension" or "freezing up on the controls" that is anathema to most flying instructors, but interpreted rather as being a natural muscular concomitant of intense application to the task.

No significant correlation was obtained between performance on the Variable co-ordination test (full hand-foot combination/balanced steering) and learning time or reaction speed on the Portable hand-foot reaction test (1); or fusion-threshold as measured on the Compact flicker-fusion machine (4); although both the latter tests have also shown a relationship to success in flying training.

Summary and conclusions

1. Full technical details are furnished of an apparatus designed for purposes of both basic research on co-ordination, and occupational selection; together with an account of its experimental application to a group of air-pilot candidates.
2. Twenty-one co-ordinative combinations can be obtained, but the present study has been restricted to the one involving both hands and both feet (balanced steering), which seemed to have the closest relevance to aircraft control.
3. "Correct Time" (i.e. "Total Time" minus "Error Time") is a more reliable measure than "Error Time," P of the difference $< 0·005$. Results obtained on two other self-paced sensory-motor tests—Steadiness (2) and Handlebars (3)—show a similar trend.
4. In the last two trials, where errors are punished by shock, rate of work is significantly slower and error less than in the preceding two trials without this stress. Irrespective of the stress element, however, normal practice effect will have contributed to the progressive improvement in error.

5. The negative correlation between "Correct Time" and "Error Time" is more substantial in the slow than the fast trials, which also fits in with findings on the above two tests, and supports the notion that the instrumental dependence between "Time" and "Error Time," with its bias towards positivity in the correlation, is more pronounced in fast than in slow working.
6. On the present test, contrary to the tendency observed on Steadiness and Handlebars, *fast* working is associated with success in flying training; but inclining towards agreement with results on Handlebars, to a small extent, less error favours the chance of success; whereas error on Steadiness showed no correlation with the flying criterion. "Correct Time" would be a better predictor of flying success than "Error Time," and of the four trials, the two last involving punishment of error by electric shock are the most significant—particularly the fourth, with a rank-difference correlation of 0·48 (P = 0·002).
7. In relation to flying success, the importance of a temperamental element in the form of fear of punishment was also noted on the Steadiness test, where the trial carrying the threat of shock had more significance than the unstressed trials.
8. Of the various points of subjective behaviour observed during the test, fear of shock, associated with failure in flying, appeared the most significant (P = 0·01). Talkativeness and sociability (P = 0·08), and the tendency to press very hard on the stirrups (P = 0·08) are both associated with success. The latter could be interpreted as an expression of strong concentration on the task.

ACKNOWLEDGEMENTS

Mr. J. B. Kirstein and Mr. R. Serdyn took the photographs. The use of the rank-difference method for test validation was suggested to me by Mr. J. Tromp and Mr. R. F. Skawran.

REFERENCES

1. DE WET, D. R. A portable hand-foot reaction test. *J. Nat. Inst. Personnel Res.*, 1958, 8, 106–116.
2. DE WET, D. R. An improved steadiness apparatus and its validity for air-pilot selection. *J. Nat. Inst. Personnel Res.*; 1959, 8, 122–136.
3. DE WET, D. R. Handlebars: a self-paced test of two-hand co-ordination, and some results on air-pilot candidates. *J. Nat. Inst. Personnel Res.*, 1960, 8, 199–208.
4. DE WET, D. R. A compact flicker-fusion machine and its application to air-pilot candidates. *Psychol. Afr.*, 1960, 9, 100–118.

CAPE & TRANSVAAL PRINTERS LIMITED, JOHANNESBURG

A SPECIFIC ARM-LEG CO-ORDINATION TEST FOR THE SELECTION OF PUPIL AIR-PILOTS

The apparatus to be described here simulates certain elements of an aircraft cockpit and provides a specific "job" test for the selection of prospective air-pilots.

The Mark 1 model was an experimental prototype designed and built by the writer during World War II in the Aptitude Tests Section of the S.A.A.F., and intended to replace an earlier Arm-leg co-ordination apparatus (SMA3) suggested by W/Cdr. G. O. Williams and evolved by Dr. E. J. Schuster of the British Medical Research Council. The SMA3, which utilized a wide variety of mechanical transmissions, was a very complicated mechanism difficult to keep standard under the rigours of long periods of sustained testing.

In designing the A.T.S. machine (Plates 1 and 2), an attempt was made to produce a simplified and less costly device which would provide a similar behavioural situation. It was hoped that fewer moving parts and those of rugged construction would make for more reliable and positive mechanical action, thereby reducing the disturbing effect of change in test standards inevitable (but not always immediately obvious) when alterations occur in the setting of the mechanism through wear, friction, strain and other causes.

The test presented a tracking task in which a drifting light-spot (corresponding to the nose of the aircraft in flight) had to be kept central on a screen by moving rudder-bar pedals left or right for lateral direction, and a "joystick" fore and aft for vertical direction. The target was a circle in the middle of the screen. When the light-spot drifted more than half its diameter across the circumference of this circle, contact systems attached to the rod carrying the light brought clocks into operation, which recorded continuously until the light-spot was steered back into the circle.

At a later stage in the testing procedure the task was complicated by the addition of a "throttle" lever which the candidate had to move backwards and forwards to extinguish two "distraction" lights situated respectively on the right and left of the stimulus panel.

Experimental applications indicated that, in general, the test performance of subjects with a hundred hours of actual flying experience was greatly superior to that of unselected recruits and while some of this difference could be attributed to acquired skill on the part of the trained men, it was unlikely that the latter would have progressed so far on the course without the requisite initial ability. On the strength of this rather rough validation, which was the best procedure possible at the time, it was therefore concluded that test scores had some potential usefulness for pilot selection.

The apparatus gave several years of service until it was superseded (some time after the war) by the Mark II model, which exacts the same tasks from the candidate.

PRESENT MODEL (MARK II)

The A.L.C. Mark II (Plates 1 to 6) incorporates the basic design of its predecessor, but achieves further precision in the detailed aspects of bearings, loci, and better proportioning of moving parts. Furthermore, where the earlier model had necessitated the use of both direct and alternating current, the present one requires laternating current only, all the recording being done on "standard" clocks which were scarce during the war period when reciprocating direct-current counters had often to serve as rather unsatisfactory substitutes.

The present apparatus was thought to have certain advantages over an electronic model of SMA3 in use by the R.A.F. and elsewhere in the Commonwealth after the war. It is mechanical in operation, requires very little maintenance, and virtually no adjustments to remain standard in performance. It is probably cheaper to construct than the electronic model, although it has a rather expensively made panelled housing, which is not strictly necessary.

Movement of the light-spot

As in the Mark I model, movement of the "drifting" light-spot is obtained very simply by a constant-speed permanent-capacitor motor driving a crank-disc to which the rod carrying the lamp is universally pivoted at its one end. At its middle the rod bears on a universal fulcrum attached to the platform carrying the motor (Plate 3). By staggering the line of the disc-axis slightly relative to this fulcrum, the actual locus described by the light spot on the screen becomes elliptical instead of circular. This "distortion" is not essential; its sole object is to make the task of keeping the light-spot central a little more difficult than it would be if the locus were a symmetrical circle. An irregular path could have been achieved by having the rod loaded against a cam of appropriate profile, instead of having it coupled at a fixed crank-distance, but the extra constructional complication was not thought worthwhile as the staggered crank system made the task sufficiently difficult.

The method whereby the manipulation of rudder-bar and stick controls the drift of the light-spot is a straightforward application of "mobilis in mobile". The platform carrying the motor and light-spot lamp is supported on horizontal bearings in a metal frame (Plates 3 and 4) which in its turn pivots vertically in a larger frame fixed to the chassis of the machine. The moving frame is coupled directly to the rudder-bar by rods underneath the chassis. Its range of movement is greater than that of the rod carrying the light-spot lamp, so that even maximum movements of the latter can be more than compensated for by suitable opposite control on the rudder-bar.

Corrections of light-spot movement in the vertical plane are accomplished by the stick, pivoting fore and aft. It is coupled directly to the platform carrying the motor through a long connecting-rod having a universal joint at either end. These joints are necessary to allow the rod to swing freely without impeding the movement of the rudder mechanism. The range of movement imparted to the platform by the stick well exceeds the maximum traverse of the crank throw.

Recording

Recording of the length of time that the light-spot is outside the limiting circle is done by means of two "wiper arms" connected universally to the lamp rod through small telescoping connecting-rods. Each of the two wipers contacts works radially across a metal plate having a triangular piece of bakelite running down the middle and recessed flush with the plate surface (Plate 6). When the wiper-contact travels beyond the limits of the bakelite on to the plate, a clock-clutch acts at the tester's control panel and the clock continues to record until the wiper is back again on the bakelite. The reason for having the bakelite piece triangular is to allow a fair margin of adjustment for sensitivity of the recording. The wiper-contact itself is adjustable along its arm. If it is set to brush across the apex of the triangular piece, recording will begin before the light-spot has touched the circumference of the limiting circle; if, on the other hand, it brushes near the base, recording will only start when the light-spot is well out of the circle. It must be set at such a position that recording will begin only when the light has travelled half its own diameter either vertically or horizontally across the circle's circumference, standardisation being on this criterion.

One set of wiper-and-plate is situated at the top of the main structure below a hatchway in the wooden housing which gives access for periodic inspection and cleaning. This set records deviations of the light-spot on the horizontal plane. The other set (Plate 5) is located on the motor platform and records deviations on the vertical plane. It is accessible from the back of the machine where two large doors are fitted to the housing (Plate 2).

Distraction Lights

The latter section of the test, in which the candidate has the additional task of extinguishing panel lights (Plate 1) as soon as they come on, by an appropriate movement of the throttle-lever, is simply an elementary choice-reaction series. Each of the two distraction light systems relays to a "Standard" clock; when a light flashes on, its clock begins to run and continues until the light is put out by moving the throttle-lever in the correct direction for that particular light, either backwards or forwards. Irregularly spaced pips on the rim of a rotating crank-disc (Plate 3) actuate two micro-switches situated tangentially to the disc, and energize the distraction light systems at intervals.

Electrical Circuits

For convenience, the tester's control-panel (Plate 1, lower) carrying the switches and clocks, etc., is designed as a separate table unit coupled by a multicord "umbilical" to the machine chassis.

The source of power is 220/230 volts A.C. stepped down through bell transformers to 6 volts A.C. for the bulb circuits of the light-spot and distraction lights and also the panel indicator. The other units, comprising clocks, relays and motor, operate directly off the mains power.

The two "Standard" clocks, recording respectively vertical and horizontal drift of the light-spot outside the circle are calibrated in minutes and seconds, whereas the two "distraction-light" clocks are in minutes and thousandths of a minutes. The latter are in effect just chronoscopes recording the candidate's choice reaction time in extinguishing the lights.

There are seven toggle-switches on the tester's panel:

(1) Mains; (2) Light-spot; (3) Driving motor; (4) Rudder; (5) Stick; (6) Clock motors (Distraction); (7) Distraction lights.

Each circuit has an appropriate safety fuse:

(a) Distraction clock clutches 1.5 A.; (b) Spot light and panel light 1 A.; (c) Driving motor ½ 0.5 A.; Rudder 0.5 A.; (d) Rudder 0.5 A.; (e) Stick 0.5 A.; (f) Clock motors (Dist.) 0.5 A.; (g) Distraction lights of 0.5 A.

All circuits are quite straightforward with the possible exception of the "distraction light" systems which incorporate two"hanging in" relays.

The throttle itself (Plate 1, upper) is simply a two-way break switch. In the upright position, which it tends to assume of its own accord by spring loading, there is continuity to each relay, but the relays are not energised until the pips on the crank-disc rim (Plate 3) actuate the micro-switches. When a micro-switch is actuated its relay closes, switching on a distraction light and starting a clock. The micro-switch contact is momentary but the relay remains closed, keeping the clock running and the light burning until the "hanging in" circuit is broken by movement of the throttle-lever backwards or forwards as the case may be.

The seating accommodation on the machine is an ordinary tubular steel chair with arm-rests (Plate 1). It is held by strap clamps and wing-nuts to the chassis so that it has some adjustment nearer to or further from the controls. In practice alteration of the average setting is only necessary in cases of very tall or very short subjects.

The illustrations help to clarify the method of operation and the mechanism of the apparatus.

PLATE 1 The general layout of controls and stimulus panel in the cockpit, and the machine in operation.

PLATE 2 The doors of the housing are open to show the mechanism from the back. The long spiral spring serves as a counterpoise to the weight of the motor and platform. Without it excessive force would have to be exercised in pushing the stick forward. Spring tension is such that when no forward pressure is put on the stick, it is balanced at about the mid position.

PLATE 3 With the counterpoise spring removed, the motor platform is viewed from above on the candidate's right. The gearing system between the motor-spindle and the crank-disc makes the latter rotate at a speed of six revolutions a minute. The rod carrying the light-spot lamp is shown extending from the crank-disc towards the glass screen through the universal fulcrum. Between the fulcrum and the lamp a universal coupling leads through the small telescoping connecting-rod to the wiper-arm of the upper recorder. On the lamp rod, between the fulcrum and the point of coupling to the crank-disc, a similar system leads horizontally to the other recorder.

PLATE 4 This is another view of the motor platform from the back on the candidate's right, only one micro-switch at the rim of the crank-disc being visible. The flat rectangular object near the tail end of the platform below the motor is the latter's permanent capacitor.

PLATE 5 Here the motor platform is seen from the back on the candidate's left, without the counterpoise spring. The "pig-tails" leading from inside the housing to the motor, micro-switch, and "wiper-and-plate" recorder, are flexible connections that allow for the up and down and sideways movements of the motor platform when the candidate works the rudder-bar and stick. The universal joint at the end of the long connecting-rod to the stick is coupled to a rigid stem extending downwards from the tail end of the motor platform.

PLATE 6 The hatchway in the top of the housing gives access to the one set of "wiper-and-plate" recorders. The knurled screw on the wiper contact is loosened to adjust it higher or lower on the arm for setting and maintaining standard sensitivity.

TESTING PROCEDURE

The test is given in two sections. First, the subject controls the light-spot with stick and rudder for a standard period; the task is then repeated with the addition of the distraction lights which have to be put out with the throttle.

When the subject has been seated comfortably with his right hand on the stick and his feet on the rudder-pedals, he is instructed as follows:

"Do you see this spot of light on the screen? If you move the stick forward (demonstration) the spot of light moves downwards, and if you move the stick backwards (demonstration) the spot of light moves upwards. Similarly, if you push the rudder with the left foot, the light will move to the left and if you push with the right foot, the light will move to the right (demonstration). Thus, by co-ordinating the movements of stick and rudder, the spot can be moved in any direction. Now move the spot of light back onto the centre of the circle on the screen.

During the test the light will tend to move by itself away from the circle. By co-ordinating the movements of the stick and rudder you must counteract the tendency of the light to wander and keep it in the centre of the circle. Should it go more than half its diameter beyond the circumference of the circle, in any direction, a clock will record against you the length of time that it is outside the circle.

Your task is therefore to keep the spot of light central, and should the light move beyond the circumference of the circle, to bring it back to the centre as quickly as possible. Is that quite clear? Now keep the light in the centre of the circle and do not move the controls until you get the starting signal. Are you ready? Begin!"

It is customary for the observer to switch on first the main supply and then the light-spot, to let the subject demonstrate for himself the effect of the controls on the light-spot. Then, when he is "set" and has the light central, "Rudder" and "Stick" switches are put on and finally, the "Driving Motor".

When the subject has performed for three minutes, the main switch is turned off and the "Error-time" in seconds on the respective "Stick" and "Rudder" clocks is recorded. Then the second section follows:

"In the next test trial, you are required to use the lever on your left as well. At irregular intervals, these two lights will switch on. To put out the right-hand light, you push the throttle back. To put out the left-hand light, you push the throttle forward. (Demonstration) It is essential to put the lights out as quickly as possible, as the time that they remain on is recorded against you. You must avoid pushing the lever backwards and forwards wildly, and it must be returned to the central position each time.

As before, the spot of light must be kept in the centre of the circle. If you allow it to drift beyond the circumference, errors will be recorded against you. Treat the lights as a side issue, but do not neglect them. Now remember, co-ordinate the stick and rudder movements, and at the same time put the lights out whenever they appear. Ready? Begin!"

This section also lasts for three minutes. The "Error-time" on the "stick" and "rudder" clocks, and the reaction-time to the distraction lights which has accumulated on the other two clocks, is recorded. The subject's ~~record~~ score is the total "Error-time" in seconds for both sections, plus the total reaction-time in the second section after it has been converted from thousandths of a minute to seconds.

Results and Discussion

As it was not practicable to do any re-testing of air-pilot candidates, an indication of the test's reliability was obtained by "split-half" correlation between the first and second sections corrected by the Spearman-Brown formula. The coefficients (Table I), computed on two separate annual groups of subjects, indicate quite good reliability, particularly in view of the fact that the second section of the test was not identical to the first because it incorporated the additional task of reacting on the throttle-lever to the distraction lights.

The relationship of test performance to success in flying training is shown in Table 2 by biserial correlations with a Pass/Fail criterion on the annual flying courses over a period of five years.

The significance levels and also the mean biserial coefficients for all the annual groups were obtained by Tate's Z*-transformation method (4, 5,). All the significance levels are two-tailled, although there would probably have been good justification for giving them as one-tailled because it was expected, in the light of previous findings with the A.L.C. Mark I, that the present test would relate to flying success.

With the exception of the 1955 group, there is a general tendency for test performance to correlate with flying success, but, because of the small size of the separate annual groups, most of the coefficients do not reach the 5% level of significance. However, the mean coefficients for the five groups, namely 0.32 (Section 1), 0.31 (Section 2) and 0.25 (Distraction lights, only), are highly significant and in the expected direction. It is likely that the validity shown by the distraction-light sub-section derives largely from its experimental involvement in the main co-ordinative task of steering the light-spot with the stick and rudder-bar. Early investigations in the Aptitude Test Section during the last war showed pretty conclusively, that this elementary kind of reaction by itself did not relate to flying success.

The validation in Table 2 was done on groups which had been highly selected for flying training by means of a battery that included the A.L.C. as one of the principal tests. This meant the inclusion of very few inferior preformances, and the resultant restriction of range in the criterion groups must have had quite a considerable effect in depressing the correlations, so that, for the most part, they are probably an under-estimate of the actual validity of the test.

Product-moment correlations between the total score on the A.L.C. test (positively accelerated) and achievement in some of the main branches of a recent flying course are given in Table 3. As the A.L.C. test was included again in the selection-battery, depression of the correlations through curtailment of the sample could also have occurred here.

As could be expected, there appears to be little or no relationship between the A.L.C. test and the "ground" or theoretical branches of the course, but the test correlates significantly (0.40) with practical airmanship. Intellectual tests such as Mental Alertness and Technical and Scientific Comprehension which correlate with certain of the theoretical branches, show no significant relationship to practical airmanship. The latter is regarded as the ultimate proof of the pilot's proficiency, and if he cannot fulfil certain minimum requirements, he does not qualify, irrespective of how well he may have fared in the theoretical examinations. It is therefore obvious that a test which shows useful validity for predicting success in practical flying, is highly desirable.

Relationships between achievement on the A.L.C. and scores on various other tests are given in Table 4. The A.L.C. appears to be a very specific test. Apart from a very small correlation with Thurstone's Repeated Letters test, it shows no connection with any of the intellectual, perceptual or temperament and personality variates. However, it shows a small to moderate overlap with Floating Effect Co-ordination (1), Variable Co-ordination (2), and the Willemse Board (3) in which a long sloping board is tilted by the subject to steer a rolling marble past obstructions into a receptacle.

Table 5 gives the mean scores on the A.L.C. and four mental tests, of subjects whose home language was respectively English or Afrikaans. Scores for the mental tests reflect positive achievement, but those for the A.L.C., the total "Error-time" in seconds. On the mental tests the English group scored significantly more. On the A.L.C., while it appears that the Afrikaans group fared rather better, the significance of the difference between the means does not reach the 5% level.

There are various possible explanations for the superiority of the English-speaking group on the mental tests which will not be discussed here apart from mentioning just two. Firstly, there is a far greater quantity and variety of technical and scientific literature available to the English-speaking in their own language; secondly, military air-pilotage as a career has less attraction for the intellectually higher-grade Afrikaans speaking than for the English-speaking of similar standard.

SUMMARY AND CONCLUSIONS

1. The mechanical functioning and application of an Arm-Leg Co-ordination test apparatus specifically designed for use in selecting candidates for air-pilot training in the South African Air Force, is described.

2. Reliability coefficients derived from "split-half" correlation between the first and second sections of the test on separate samples of 161 and 185 subjects, were respectively, 0.89 or 0.91.

3. On five consecutive annual groups totalling 181 pupil pilots, The following mean biserial correlations were obtained between achievement on the test and success or failure in flying training: Section 1, Stick and Rudder Co-ordination,

0.32 (P = 0.0002); Section 2, Stick and Rudder Co-ordination plus Throttle-light distraction, 0.31 (P = 0.0003); Throttle-light distraction by itself, 0.25 (P = 0.0042).

4. A product-moment correlation of 0.40, significant at better than the 5% level was obtained between total performance on the A.L.C. and achievement in practical airmanship.

5. The above validity coefficients probably give under-estimates of actual validity because of restriction of range in the criterion groups through pre-selection.

6. There appears to be no significant relationship between the A.L.C. and the theoretical branches of the flying course.

7. The A.L.C. is a very specific test, but it does show small to moderate correlations with three other co-ordinations tests.

8. While subjects whose home language was English tended to do better at certain mental tests than those whose home language was Afrikaans, there was no significant difference in the mean performances of the groups on the A.L.C.

9. Based essentially on an earlier prototype devised by the writer in the Aptitude Tests Section of the S.A.A.F. during the last war, the present model of the A.L.C. incorporates certain mechanical improvements and during the past fifteen years it has been a standard test in all S.A.A.F. pilot selection batteries. Shortly after the machine had been completed and tried out, full constructional details were

made available to the Royal Air Force in Britain. It has been reported that the R.A.F. constructed and are using, a similar device they call the "Springbok".

ACKNOWLEDGEMENTS

Construction of the Mark II model was done in the Central Workshops of the South African Council for Scientific and Industrial Research under the supervision of Mr. J. van der Staay. Mr. S. Harrison did the mechanical work; Mr. R.J. Botes, the woodwork; and Mr. S.J. Botes the electrical installation.

REFERENCES

1. DE WET, D.R. Co-ordination and floating effect
Nat. Inst. Personnel Res., 1959, 8, 28-38.

2. DE WET, D.R. A variable co-ordination test and its potentiality as a gauge of aptitude for airmanship. Psychol. Afr., 1962, 9, 86-99.

3. SKAWRAN, R.F. The Willemse board as a personality test. Psychol. Afr., 1962, 9, 72-85

4. TATE, R.F. The biserial and point correlation coefficients. Univ. N. Carolina Inst. Math. Stat. Mimeographed Series No. 14.

5. WALKER, H.P. and J. LEV. Statistical Inference New York, Henry Holt, 1953. (refers to Tate above 269-270)

TABLE 1

Correlation between first and second section of test.

Group	N	r	Rel.
1	161	0.80	0.89
2	185	0.84	0.91

TABLE 2

Biserial correlations with Pass/Fail Criterion.

Year	1952	1953	1954	1955	1956	Overall
N	29	23	38	48	43	181
Passes	25	18	30	34	33	140
Failures	4	5	8	14	10	41
Co-ord. Section						
r bis	0.38	0.27	0.43	0.02	0.52	0.32
P	0.0930	0.1892	0.0182	0.8886	0.0019	0.0002
Co-ord. + Distraction						
r bis	0.31	0.35	0.38	-0.02	0.63	0.31
P	0.1586	0.1616	0.0536	0.8886	0.0002	0.0003
Distraction Lights only						
r bis	0.17	0.19	0.32	0.26	0.30	0.25
P	0.4778	0.4716	0.0930	0.1470	0.0970	0.0042

TABLE 3

Relationship of A.L.C. and two mental tests to various Flying training criteria.

(Decimal points omitted)

N	Flying Training Criteria	A.L.C. Total	Ment. Alert.	Tech. S. Comp.
55	Aero Dynamics	17	18	31s
57	Meteorology	-18	33s	31s
55	Aviation	18	-04	18
55	Engines	-23	11	09
55	Instruments	22	15	31s
63	Radar	02	06	25s
55	Approach Aids	06	39s	27s
56	Navigation	15	28s	23
54	Ground %	05	26s	36s
55	Flying %	40s	04	10

"S" implies significance at the 5% level or better.

TABLE 4

Correlations between A.L.C. and other tests.
(Decimal points omitted)

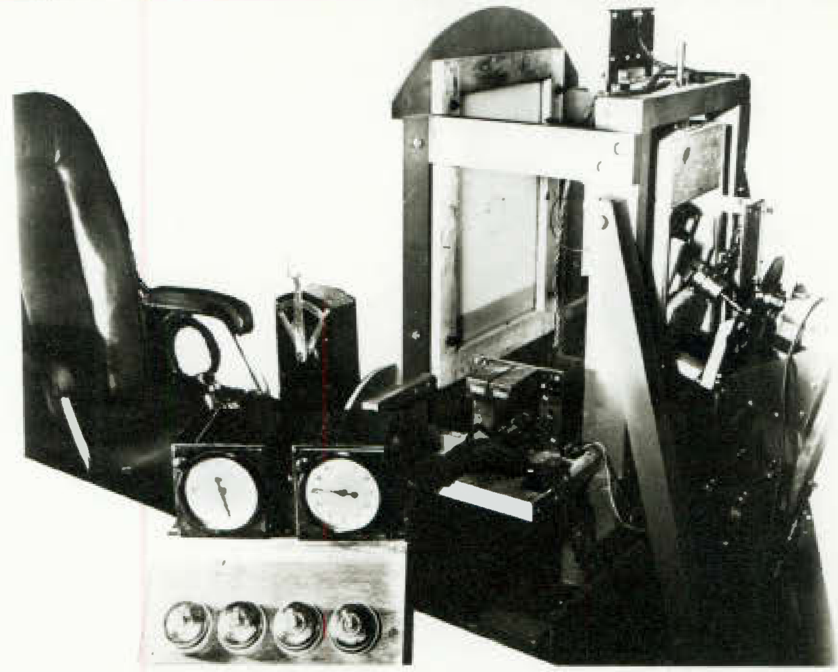
N	Other Tests	A.L.C.
177	Mental Alertness	01
177	Technical and Scientific Information	-09
177	Technical and Scientific Comprehension	-01
177	Flying Interest Inventory	-04
176	Mathematics	05
132	Gottschaldt Figures	05
132	Designs Test	-09
120	Dots Cancellation	10
132	Repeated Letters	22s
177	Span of Attention	11
177	Serial Discrimination	-07
177	Flicker Fusion	-02
113	E.E.G. Alpha Frequency	06
128	Personality Inventory	02
172	Temperament Questionnaire	09
177	Steadiness (Manual)	12
177	Floating Effect Co-ordination	25s
177	Variable Co-ordination	26s
177	Willemse Board (Co-ord.)	32s

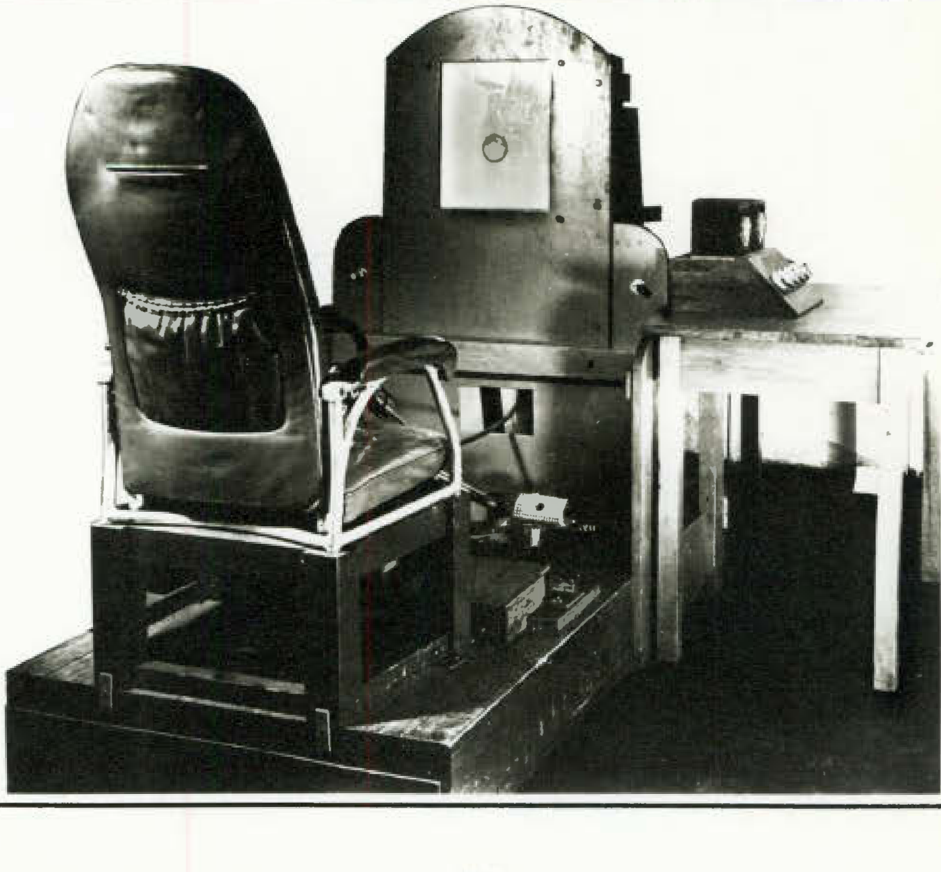
TABLE 5

Means and S.D.s of English and Afrikaans speaking subjects.

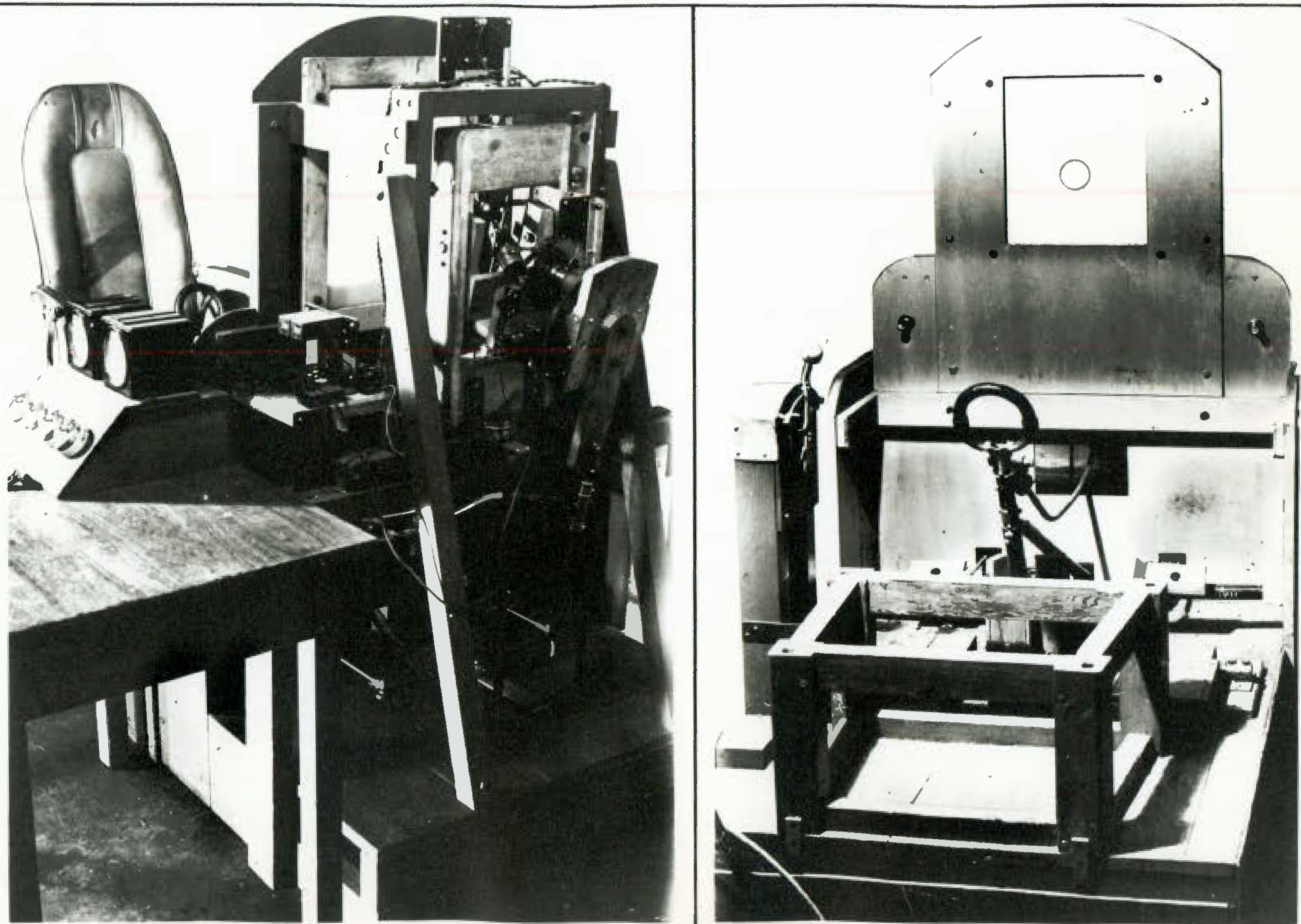
TEST	English N=72		Afrikaans N=105		C.R. of Diff. M's
	Mean	S.D.	Mean	S.D.	
Tech. and Scient. Information	37.49	4.46	29.99	5.44	9.98
Tech. and Scient. Comp.	22.26	3.78	19.77	4.27	4.06
Flying Interest Inventory	54.67	12.38	45.80	10.35	4.96
Mental Alertness	27.76	3.98	24.49	5.27	4.68
Arm-leg Co-ordinator	379.19	74.13	364.96	85.11	1.17

AIR-PILOTS' ARM-LEG CO-ORDINATOR MARK I. PLATE 1.

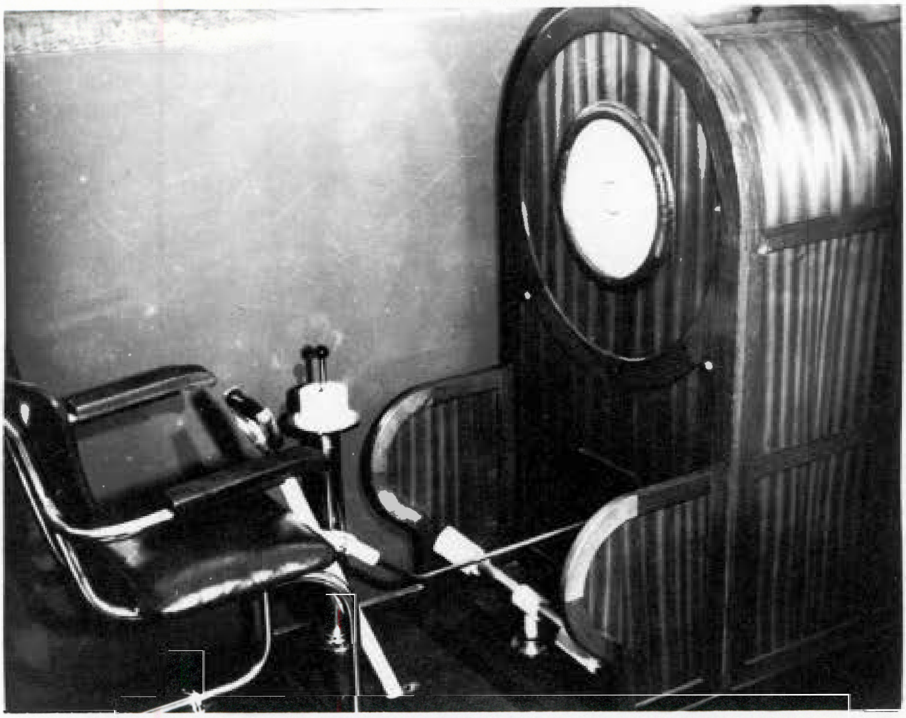




AIR-PILOTS' ARM-LEG CO-ORDINATOR MARK I. PLATE 2.

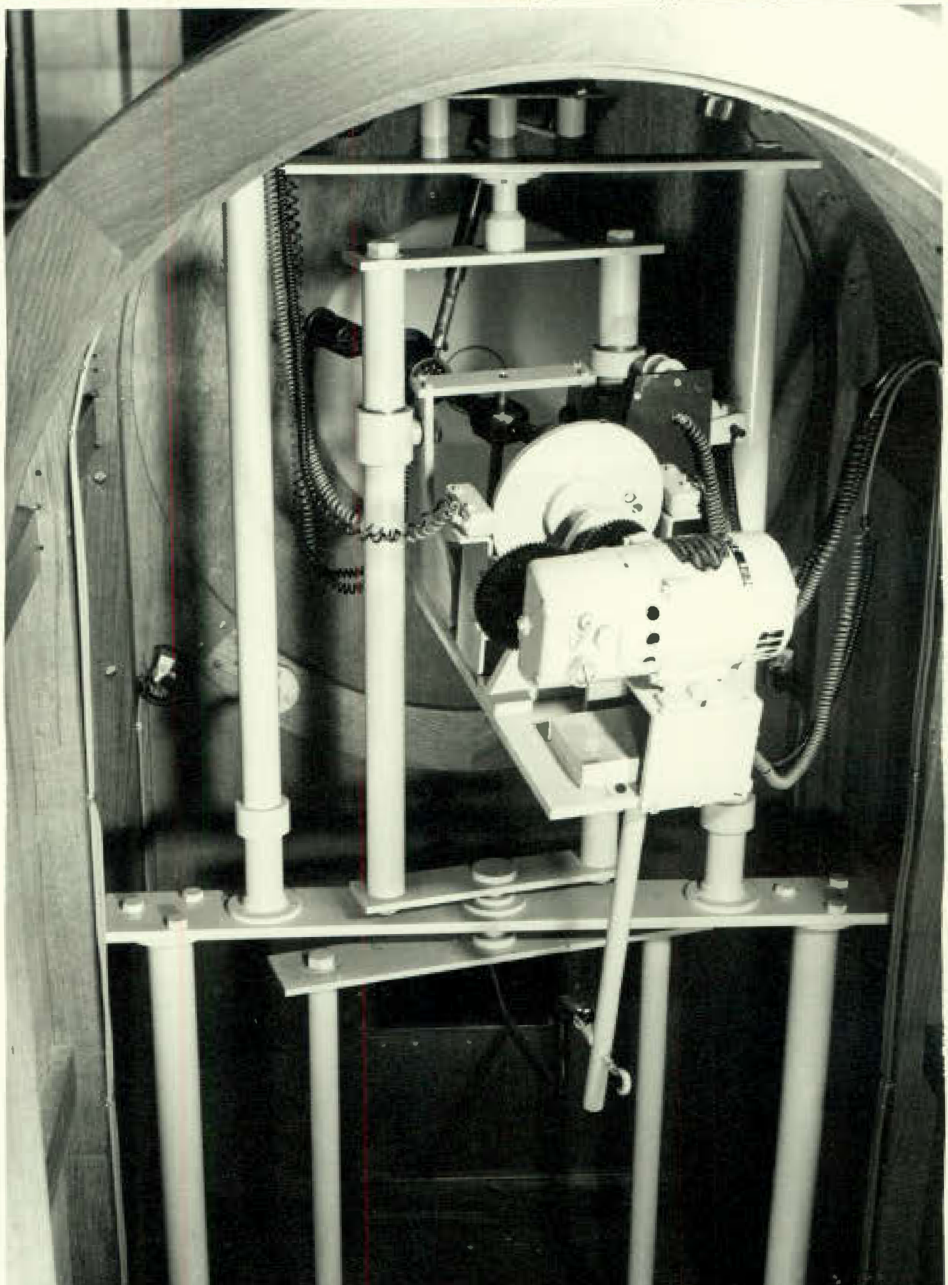


AIR-PILOTS' ARM-LEG CO-ORDINATOR MARK II. PLATE 1.

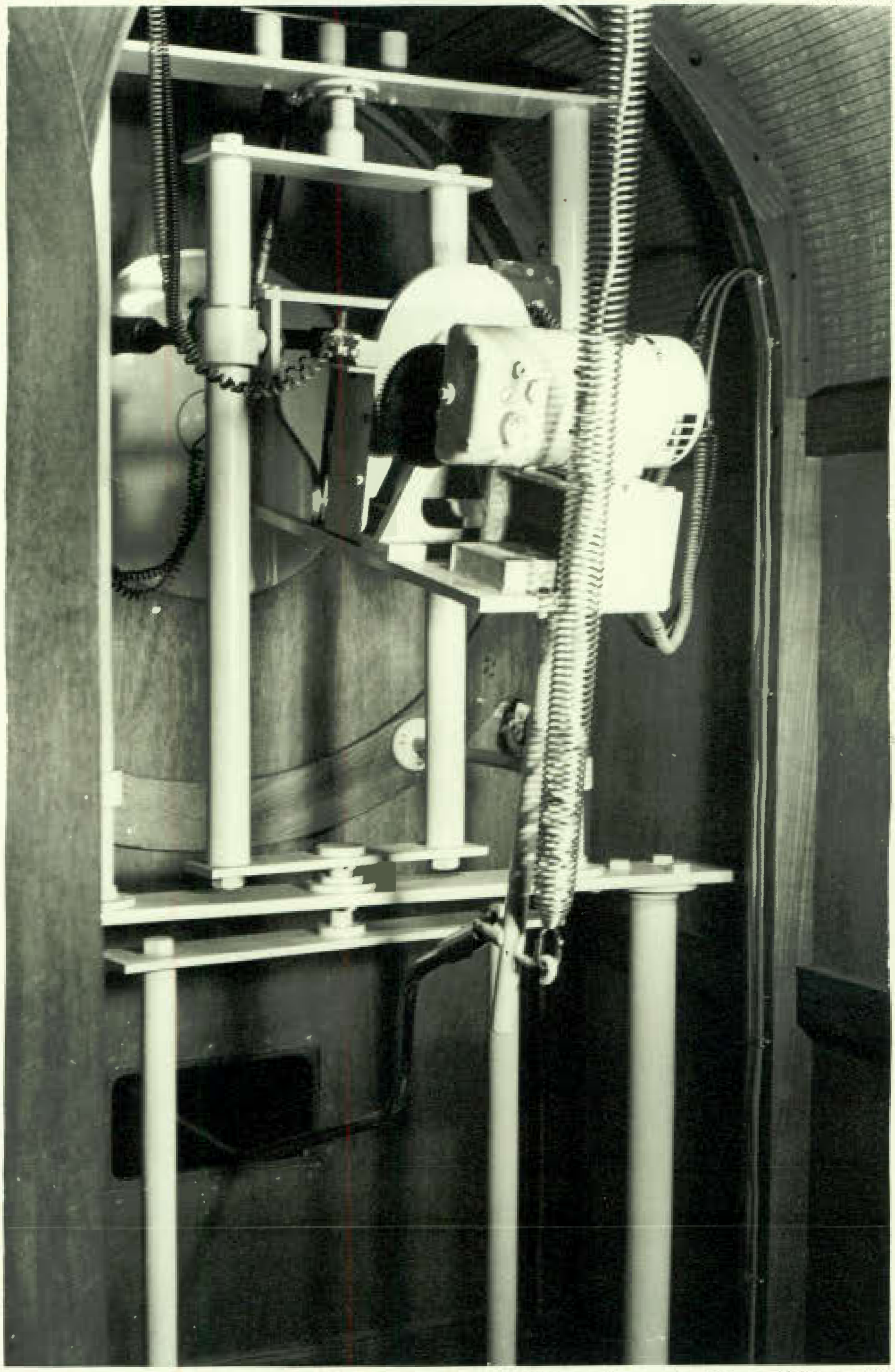


AIR-PILOTS' ARM-LEG CO-ORDINATOR MARK II. PLATE 2.

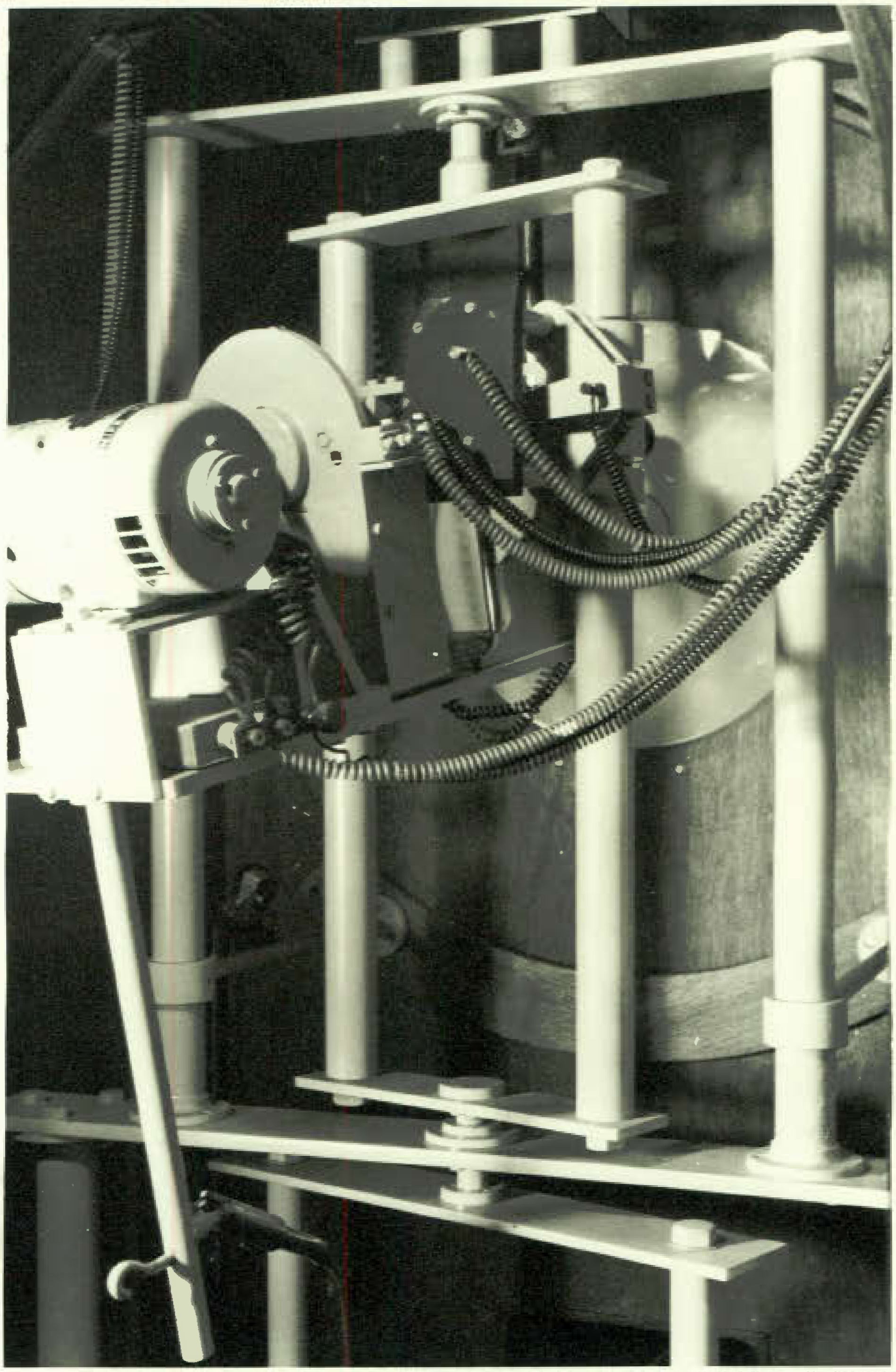
AIR-PILOTS' ARM-LEG CO-ORDINATOR MARK II. PLATE 3.



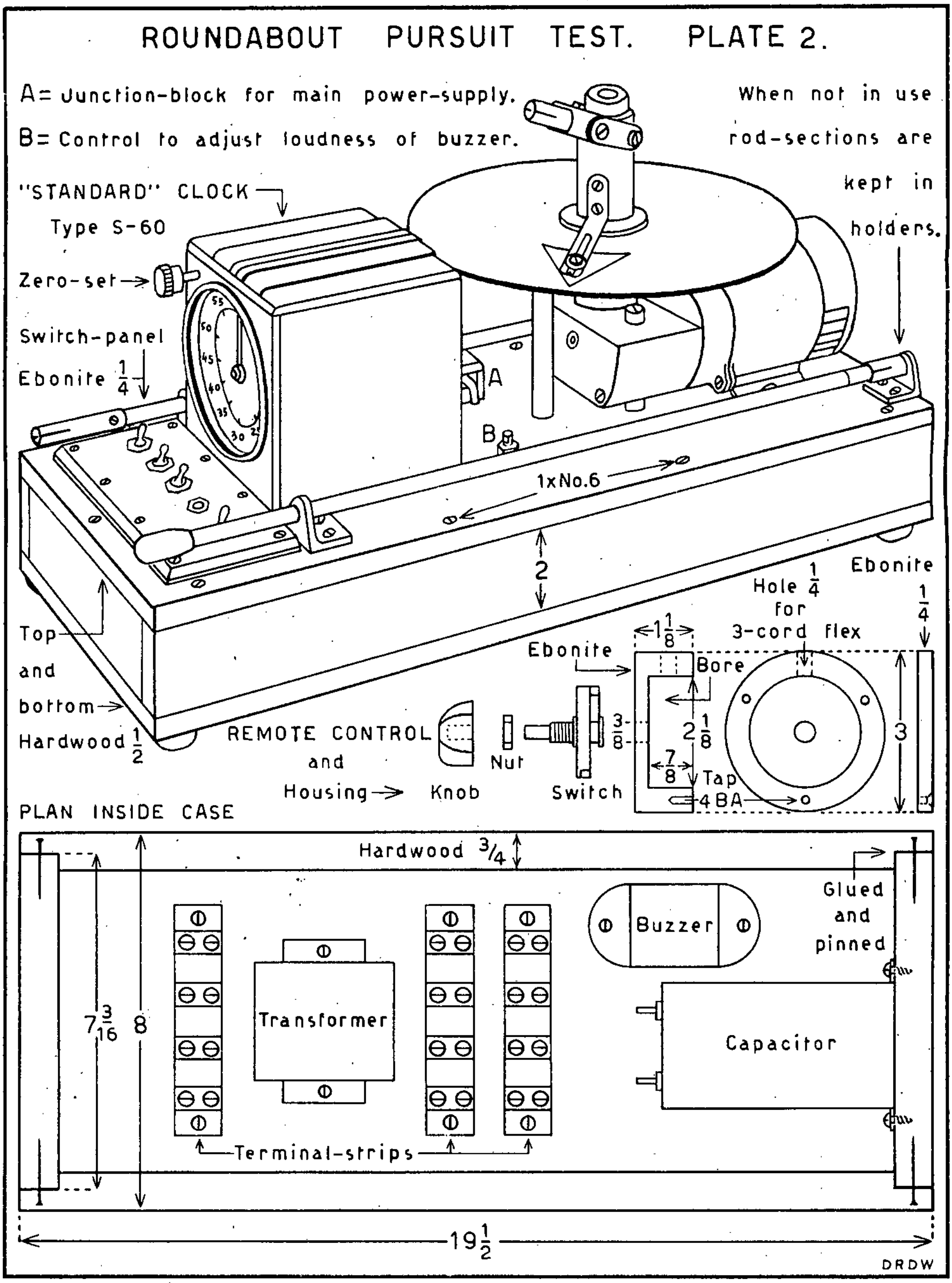
AIR-PILOTS' ARM-LEG CO-ORDINATOR MARK II. PLATE 4.



AIR-PILOTS ARM-LEG CO-ORDINATOR MARK II. PLATE 5.



AIR-PILOTS' ARM-LEG CO-ORDINATOR MARK II. PLATE 6.


ROUNDABOUT PURSUIT TEST. PLATE 2.
A= Junction-block for main power-supply.
B= Control to adjust loudness of buzzer.
When not in use rod-sections are kept in holders.
"STANDARD" CLOCK
Type S-60
Zero-set
Switch-panel
Ebonite 1/4
A
B
1xNo.6
2
Top and bottom Hardwood 1/2
Ebonite
Hole 1/4 for 3-cord flex
Ebonite 1/4
1 1/8
Bore
3/8
2 1/8
3
7/8
Tap
4 BA
REMOTE CONTROL and Housing
Knob
Nut
Switch
PLAN INSIDE CASE
Hardwood 3/4
Glued and pinned
Buzzer
Transformer
Capacitor
7 3/16
8
Terminal-strips
19 1/2
DRDW

which on the smooth metal surface proved too difficult when attempted simultaneously with the other task.

The pursuit-unit is a straightforward piece of electro-mechanical equipment made up as follows. A rectangular wooden case constitutes both a housing for certain electrical components, and a platform for mounting a switch-panel, a clock, a constant-speed motor driving a metal disc, and some smaller fittings, including holders for the two sections of the tiller-rod.

Pivoting on an extension of the motor-spindle, but insulated electrically from the disc, is a vertical bush having at its upper end a fork and thimble to provide linkage with the tiller-rod, and at its lower, a pointer projecting out radially over the disc. This pointer is slotted to allow of adjustment in the position of a small stud that clamps to it and makes light spring-loaded contact with the disc-surface.

On the disc, extending back from its rim, an inset of bakelite forms a non-conductive area between disc and stud, and being triangular with its apex outwards, permits the difficulty of the subject's task to be pre-set at a higher or lower level by clamping the stud at a position on the pointer nearer to or further from the periphery of the disc. (For all testing done in connection with the present study, the stud was kept clamped at a constant position seven-eighths of an inch inwards from the apex of the inset.)

As shown in Plate 4, a spring-loaded brush at the top of a vertical pillar under the disc connects the latter to a wire leading to the clock-clutch and a potentiometer that functions simply as a convenient means of controlling the loudness of the buzzer with which it is in series. Another wire couples one 6-volt pole on the transformer with the motor-housing, and continuity extends via the motor-spindle to the pointer and stud, but only includes the disc when the stud is off the bakelite inset and in contact with the metal surface. For the duration of such contact the buzzer sounds and the clock records.

Switch A controls the main power-supply; B enables the clock and buzzer to be isolated when the disc is run on to the standard starting position for each section of the test by means of D. The transfer-switch C potentializes either D for the above purpose, or the remote-control switch E used during the test for stopping and starting the motor and changing its direction of rotation. E is mounted in a small ebonite case and has a long three-cord cable terminating in a jack-plug that couples to a socket in the switch-panel. Rotation of the disc in either direction is always at a constant speed of four revolutions per minute.

Two accessories that greatly facilitate operation of the remote-control in a standard manner are the index-rings for the observer's stop-watch, shown in Plates 1 and 3. Constructionally identical, each has on its flange a particular series of painted marks to indicate the time-intervals for changes in control. The ring fits snugly in position around the stop-watch dial, and with the test-trial in progress, as the stop-watch hand passes each mark in turn, the observer gives an appropriate turn on the knob to stop and start the disc or change its direction of rotation.

Testing procedure

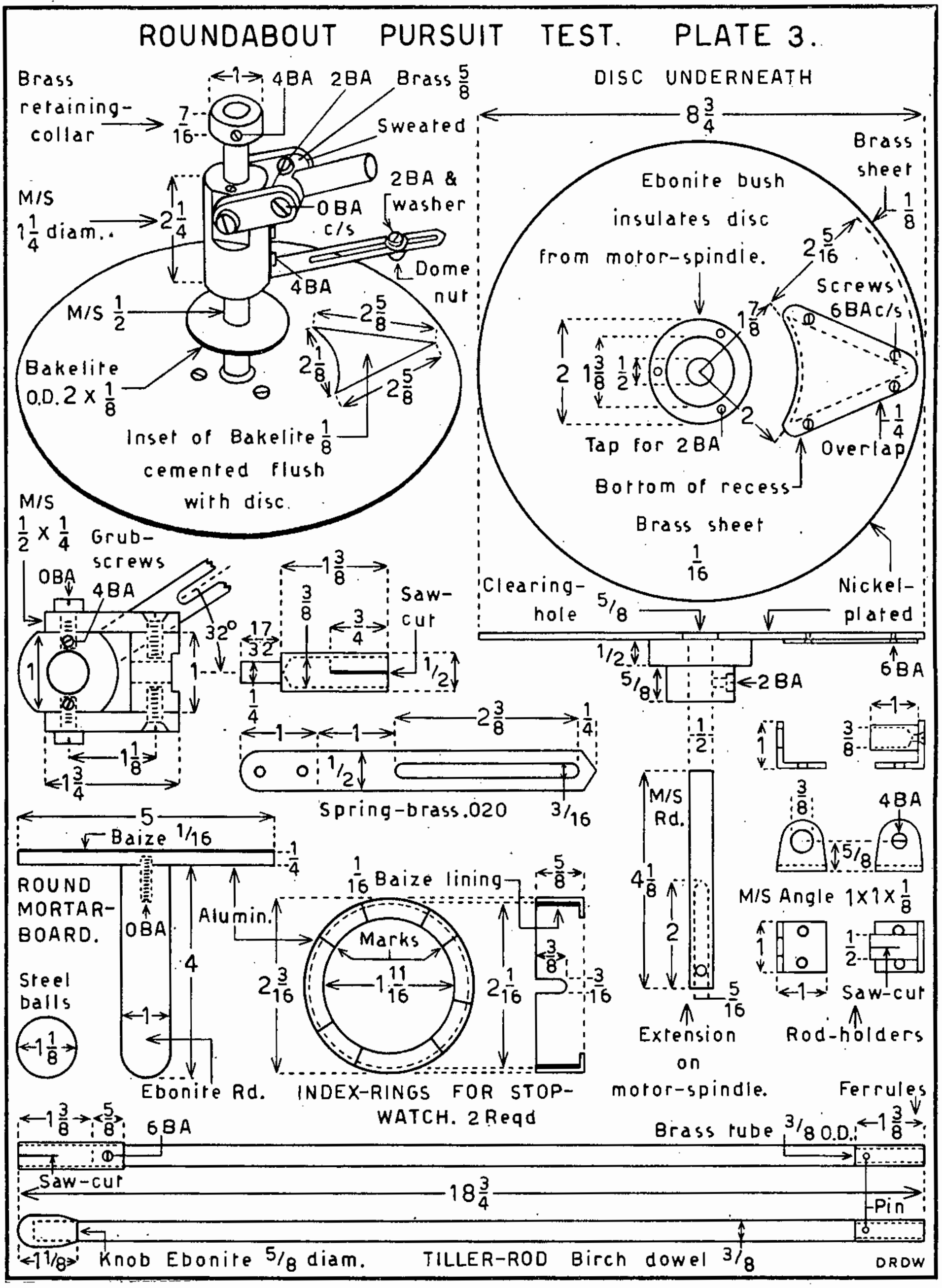
The upper photo in Plate 1 shows how the apparatus is arranged. The pursuit-unit rests on a small square table in the middle of the room, its disc being more or less central with the table-top. Two trays, each containing six steel balls, are so placed that one or other will be easily accessible to the subject from any point round the table. The remote-control switch is fixed with a light G-clamp to the observer's table that stands well to one side, and the cords leading from both this control and the main-supply to the pursuit-unit are covered with a piece of linoleum to prevent the subject from tripping over them. Having assembled the tiller-rod in its fork and set the bakelite mark on the disc at the standard starting position (i.e. pointing towards the corner of the room diagonally opposite from where the observer will sit), the observer demonstrates to the subject what is required.

" The disc with the black mark will be turning steadily, and by means of the rod which you hold in your left hand at the knob, you must try to keep this pointer over the black mark. Whenever you allow the pointer and mark to drift apart, a buzzer will sound, which means that error is being recorded against you, so get them together again as soon as possible. Now while you are doing this and moving round the table, you must keep a steel ball balanced like this with your right hand. Both tasks are equally important; the longer you keep the ball balanced and the more carefully you keep the pointer on the black mark, the better your score will be. If the ball rolls off, don't try to pick it up; just take another from one of these trays for the next trial. Get into position now and see that the pointer is on black and the ball nicely balanced."

At the remote-control, the observer says, " Ready? Go! " and sets the disc rotating forwards, at the same time starting the stop-watch. When the subject lets the ball drop, both watch and disc are stopped, and the duration of the trial is recorded on the scoring-sheet. However, should the ball be kept on the mortar-board, the trial is prolonged for thirty-seconds and this figure recorded.

Using either the same ball, or taking another from a tray, the subject gets it balanced in readiness to start the second trial from the position where the first ended. The observer says, " Ready? Go! " and as the stop-watch is started, sets the disc going in reverse. This trial ends likewise when the ball drops or the maximum duration of thirty seconds is attained, and the time is recorded as before. For the third trial, the subject takes another ball from a tray if necessary, and carries on from where he stopped in the second, but the observer sets the disc running forwards again. Procedure in the fourth trial is identical, except that the disc runs in reverse. These four trials comprise Section A (Plain Circuits), and the subject is allowed a brief rest while the observer records the " Error Time " on the clock (i.e. total period during which the pointer and disc-mark were out of alignment), resets the clock to zero, and runs the disc on to the starting point with the mark pointing towards the corner of the room. The subject is then told: " You start another series now; take up your position with the ball balanced and the pointer in line."

The observer slips the index-ring labelled B over the stop-watch dial, says " Ready? Go! " and starts the disc rotating forwards. As the stop-watch hand passes each mark on the ring, the control-switch is turned to alter direction of rotation, and in this series also, activity continues for thirty seconds or ends when the ball drops. The next trial is carried on from where the previous one ended, rotation starting in reverse this time and the direction alternated accordingly as the stop-watch hand passes the marks on the index-ring. A third trial, starting forwards again, completes Section B (Simple Alternations) and the subject gets


ROUNDABOUT PURSUIT TEST. PLATE 3.
Brass retaining-collar
4BA 2BA Brass 5/8
Sweated
M/S 1 1/4 diam.
2BA & washer
0BA c/s
Dome nut
4BA
M/S 1/2
Bakelite O.D. 2 x 1/8
Inset of Bakelite 1/8 cemented flush with disc.
DISC UNDERNEATH
Brass sheet
Ebonite bush insulates disc from motor-spindle.
Screws 6BA c/s
Tap for 2BA
Overlap
Bottom of recess
Brass sheet 1/16
Clearing-hole 5/8
Nickel-plated
6BA
2BA
M/S 1/2 x 1/4
Grub-screws
0BA
4BA
Saw-cut
Spring-brass .020
M/S Rd.
4BA
M/S Angle 1 x 1 x 1/8
Baize 1/16
ROUND MORTAR-BOARD.
Alumin.
0BA
Steel balls
Ebonite Rd.
1/16 Baize lining
Marks
INDEX-RINGS FOR STOP-WATCH. 2 Reqd
Extension on motor-spindle.
Saw-cut
Rod-holders
Ferrules
6BA
Brass tube 3/8 O.D.
Saw-cut
18 3/4
Pin
Knob Ebonite 5/8 diam.
TILLER-ROD Birch dowel 3/8
DRDW

another short rest while the observer notes the " Error Time " on the clock, resets to zero, and runs the disc back to the starting point for Section C (Complex Alternations).

The index-ring labelled C, used on the stop-watch in this last series, has some marks for imposing changes in direction of rotation, and others (bracketed in the note below) to provide a momentary stopping of the disc and resumption of movement in the same direction. Procedure is essentially the same as for Section B, three trials being given, starting respectively: reverse, forwards, reverse. As before, the duration of each trial up to the maximum of thirty seconds, and also the " Error Time " that has accumulated on the clock, is recorded.

The calibrations on the index-rings correspond to the following second-intervals on the stop-watch:

Ring B: 3, 6, 10, 12, 16, 19, 24, 27.

Ring C: 4, (7·5, 8·5), 11, 14, (15·5, 16·5), 21, (23·5, 24·5), 27.

The above procedure, which was used in the experimental application of the test to be described here, has a distinct weakness that should be remedied in any future work. This concerns the ending of a trial when the subjects let the ball drop, which benefits those who are unable to keep the ball balanced for long, in so far as they have less time in which to accumulate error. It would be much better to have an equal total amount of actual balancing time for each subject. Recording would be stopped temporarily when a ball dropped, and then resumed when it had been replaced. In this latter method the two measures would be the total period during which the pointer was off the mark on the disc, and the number of balls dropped; total time spent in actual performance being an equal quantity for all subjects, although the number of breaks in the continuity varied from one subject to another.

Application and results

The test was applied experimentally in 1956 to a group of air-pilot candidates for the South African Air Force, with the main object of ascertaining its usefulness as a predictor of success or failure on a course of flying training, but also to obtain some information on the inter-relationship between measures and sections within the test and the extent of its relationship to certain other sensory-motor tests.

Selection of the forty-seven candidates (Table 1) who were accepted for training was done on the basis of other testing and interview procedures, and performance on the Roundabout carried no weight. For validation purposes, five candidates who were suspended from the training course for reasons other than poor airmanship, have been excluded.

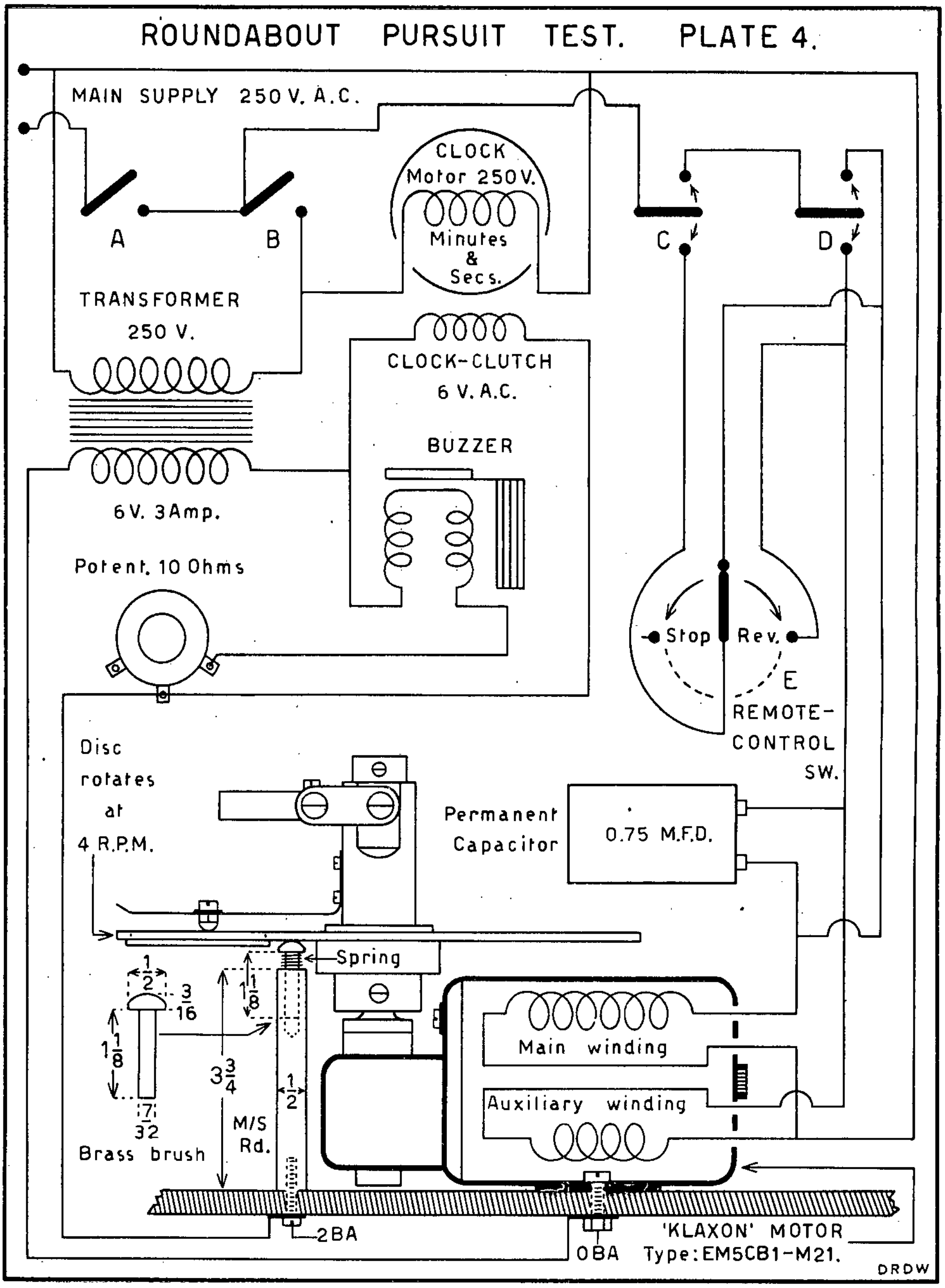

ROUNDABOUT PURSUIT TEST. PLATE 4.
MAIN SUPPLY 250 V. A.C.
A
B
TRANSFORMER
250 V.
6 V. 3 Amp.
CLOCK
Motor 250 V.
Minutes
&
Secs.
CLOCK-CLUTCH
6 V. A.C.
BUZZER
C
D
Potent. 10 Ohms
Stop
Rev.
E
REMOTE-
CONTROL
SW.
Disc
rotates
at
4 R.P.M.
Permanent
Capacitor
0.75 M.F.D.
Spring
Main winding
Auxiliary winding
Brass brush
M/S
Rd.
2BA
0BA
'KLAXON' MOTOR
Type: EM5CB1-M21.
DRDW

Table 1

Derivation of sample

Year	1956
[Applied for flying course	163
[Did Roundabout test	161
Accepted for training	47
Suspended from training (all reasons)	13
[Suspended for poor flying ability	8
[Did Roundabout test	8
[Completed course successfully	33
[Did Roundabout test	32

As Section A of the test consists of four trials, and B and C of three each, to make direct comparison easier the means in Table 2 are given in terms of the average time per trial for each section.

Table 2

Means in terms of the average time per trial in seconds for each of the three sections of the test

Number of cases = 161			
Three sections of test	A	B	C
Number of trials in each ..	4	3	3
Possible duration per trial ..	30	30	30
Average duration achieved ..	16·8	13·0	15·0
Average " Error Time ".. ..	2·5	3·0	3·3
S.D.	1·8	2·0	2·0
Average " Correct Time " ..	14·3	10·0	11·7
S.D.	7·5	6·0	6·7
Critical Ratios of the differences			
Sections of test	A & B	A & C	B & C
Average " Error Time ".. ..	4·10	6·56	2·24
Average " Correct Time " ..	8·32	4·86	3·96
($t \geqslant 2{\cdot}24$ implies significance at better than the 3% level)			

Two opposing influences are present; firstly, the effect of practice which makes for improvement, and secondly, some increase in basic difficulty between the sections that retards improvement. In Section A, successful activity is maintained for a longer period and less error made than in B and C. Between A and B there is a significant reduction in the period of successful activity and an increase in error-time, together reflecting the substantially greater basic difficulty of Simple Alternations over Plain Circuits. In C there is a marked *increase* in the period of successful activity over that in B, accompanied by an increase in error-time considerably less significant than it was between A and B. Present results do not warrant any conclusion on whether the gradient of basic difficulty between Plain Circuits and Simple Alternations is greater or less than that between Simple Alternations and Complex Alternations, but in either case it seems likely that the cumulative effect of practice would have been greater at the later stage of the test.

Intercorrelations between the three sections of the test are given in Table 3. These values will have been depressed because the introduction of changes of direction in the second section and stops in the third made the three sections rather different, and in a sense, separate tests. Nevertheless they can be used to give an indication to the lower bound of the reliability of the procedure as a whole. By applying the Spearman-Brown prediction formula, using the average of the inter-section correlations as the lower bound estimate of reliability of a single section of the test, and regarding all sections as being of approximately equal length, then the overall reliability would be of the order of 0·80 for " Correct Time " and 0·84 for " Error Time."

Table 3

Product-moment intercorrelations between sections on the measures " Correct Time " and " Error Time "

N = 161	Correct Time	Error Time
A & B A & C B & C	0·55 0·55 0·64	0·64 0·63 0·62
Mean r P of Mean r	0·58 <0·0001	0·63 <0·0001
P of the difference between Mean r's >0·45		

The correlations between " Correct Time " and " Error Time " in Table 4 are practically zero for Section A, but significant at the 10 per cent level for B, and at better than the 0·01 per cent level for C. In the Roundabout test, even although instructions stress the equal importance of both activities, the subjects cannot be compelled to distribute their attention in an identical and consistent manner. Differences are likely, both between the performances of individuals and between stages in the individual's own performance, and the latter differences in particular, could be expected to occur most markedly during the early section of the test where the subject finds himself in a strange situation and under the necessity to explore for some suitable system whereby he can accomplish two exacting tasks simultaneously. The instability characterizing this experimental period seems a probable reason for the apparent lack of any relationship between the two activities in Section A.

Table 4

Product-moment correlations between " Correct Time " and " Error Time " for each of the three sections

N	161			P of the differe nces betwee left		
Sect. r P	A −0·05 0·529	B 0·13 0·099	C 0·23 0·004	A & B 0·097	A & C 0·009	B & C 0·340

With regard to results obtained by the author on Handlebars (1) and Variable Co-ordination (2), it was mentioned that in self-paced sensory-motor tasks there is a factor of instrumental dependence that makes for some positivity in the correlation between " Error Time " and " Time." This implies that the longer the duration of the activity, the greater the scope for the accumulation of " Error Time," and to make such a relationship possible, it is of course necessary that the duration of the activity should also be a test-variate. This requirement is met in the case of self-paced tasks affording measures of both quantity and quality, like the above two tests, but not in certain " pursuit " tasks exemplified by the Koerth rotor (3) and the Carl Heinrich co-ordination apparatus (Melton, 4) where duration of activity is a constant and there is only a qualitative measure of time on or off a criterion. In the latter category, therefore, the question of such dependence does not arise.

Although the Roundabout is also a " pursuit " test, in so far as a constant rate of work is imposed on all subjects, the fact that both the duration of the activity and its quality are test-variates, makes it, in another respect, more closely identifiable with the self-paced type, and the presence of some instrumental dependence between the variates can be reasonably assumed. This would be supported by the significant positive correlations between " Correct Time " and " Error Time " in Sections B and C (Table 4). It is possible, that in Section A where there is no correlation, the effect of this dependence has been overrun by stronger influences relating to adaptability in the new situation.

Biserial correlations between performance on the Roundabout and subsequent achievement in flying training are given in Table 5. " Correct Time," which reflects the duration of successful activity in keeping the ball on the mortar-board, shows no significant validity. On the other hand, the aggregate of " Error Time," which is a measure of failure in keeping the pointer synchronized with the rotating disc, has a correlation of — 0·35, significant at better than the 7 per cent level. Whether subjects in general tended to allocate more attention to the rotating disc than to the ball and mortar-board, cannot be inferred from the present data. With regard to the three sections A, B and C of the test, the differences between the respective correlations are not significant, and there is accordingly no indication that any one section would be more useful than another for the prediction of flying ability.

Table 5

Biserial correlations between performance on the Roundabout test and success/suspension in flying

Section of test	Correct Time		Error Time	
	r_{bis}	P	r_{bis}	P
A. Plian Circuits	0·20	0·312	−0·25	0·184
B. Simple Alternation	0·02	0·897	−0·33	0·078
C. Complex Alternations	−0·17	0·412	−0·30	0·114
Aggregates	0·05	0·749	−0·35	0·066
N = 40 (Successes 32, Suspensions 8)				

To sound the extent to which censoring of the distribution as shown by the selection ratio in Table 1 would have affected the validities obtained, two biserial correlations (not given in the tables) were done on Roundabout performance between the group accepted to undergo flying training and the rest of the total sample who were rejected. On aggregate " Correct Time " the biserial was zero; on aggregate " Error Time " it came to —0·18 (P = 0·063); the number of cases being 161. Censoring has probably had a slight adverse effect on the validity of the " Error Time " measure.

In Table 6, showing the correlations between the Roundabout and certain other sensory-motor tests, a point of general interest is the similarity in pattern between the columns for " Total Time " and " Correct Time," produced by the obvious close connection between these measures.

Table 6

Product-moment correlations between the Roundabout and other sensory-motor tests

	ROUNDABOUT					
	Total Time		Error Time		Correct Time	
	r	P	r	P	r	P
FLICKER CFF (Aggregate)	—0·11	0·165	—0·01	—	—0·12	0·129
TWO-HAND CO-ORDINATION (Moede Type)						
Errors	0·06	—	0·00	—	0·06	—
Time	—0·30	<0·0002	0·16	0·043	—0·35	<0·0001
ARM-LEG CO-ORDINATION (Air-pilots)						
Without distraction	—0·17	0·032	0·09	—	—0·20	0·011
With distraction	—0·25	0·001	0·00	—	—0·27	<0·001
STEADINESS (Manual)						
Total Time	—0·02	—	—0·01	—	—0·02	—
Error Time	—0·19	0·016	0·03	—	—0·20	0·011
HAND-FOOT REACTION						
Learning Time	—0·07	—	0·02	—	—0·08	—
Speed	—0·01	—	—0·15	0·057	0·03	—
Errors	—0·06	—	—0·07	—	—0·04	—
HANDLEBARS CO-ORDINATION						
Total Time	—0·09	—	0·01	—	—0·09	—
Error Time	—0·25	0·001	—0·07	—	—0·24	0·002
	Number of cases = 161					

" Error Time " correlates to a small but significant extent with speed on Two-hand co-ordination (Moede type) and Hand-foot reaction, which means that there is a slight tendency for fast workers on these tests to do better than slow ones in keeping the pointer synchronized with the disc on the Roundabout.

All the more substantial correlations are on Roundabout " Time " (Total or Correct), and the largest of these is with " Time " on Two-hand co-ordination (Moede type), implying that proficiency in keeping the ball balanced is linked with fast two-hand co-ordination. In the Arm-leg co-ordination test subjects had to use an aircraft-type stick and pedal-bar to keep a drifting lightspot on a target. The second section of this test, with the additional complication of distraction-lights that had to be extinguished by reacting on a " throttle " lever, appears to relate more significantly to " Time " on the Roundabout than the first section, probably because of the increased demand for distribution of attention. Success in keeping the light on the target is associated with success in keeping the ball balanced.

" Total Time " on the Handlebars test has no connection with any measure on the Roundabout, but " Error Time " correlates with " Time " on the Roundabout, which denotes that careful work on Handlebars is associated with prolongation of the balancing-activity on the Roundabout.

The correlation between " Time " on the Roundabout and Flicker-fusion threshold misses acceptable significance, but does not run counter to the hypothesis that the less inert subjects with a higher fusion-threshold should have a rather better faculty for distribution of attention and be less disorganized by a sudden change of set than the more inert or " secondary " subjects.

From the practical standpoint, the most noteworthy fact emerging from this table is that " Error Time " on the Roundabout relates much less to the other tests than " Correct Time," whereas it surpasses " Correct Time " in having some validity for flying training, and this suggests that " Error Time " on the Roundabout might be a useful contribution to a battery of sedentary tests. " Error Time " being the measure intrinsic to the pursuit-unit that actually keeps the subjects moving, is perhaps more directly representative than " Correct Time " of those factors of overall bodily agility and kinaesthetic control which are lacking in the sedentary tests.

Summary and conclusions

1. This paper describes the construction and experimental application of a sensory-motor test involving postural balance, co-ordinated response from diverse body-musculature, distribution of attention, and flexibility in reacting to a sudden change of " set."
2. The test comprises the two-fold task of balancing a ball on a mortar-board while moving about to keep a pointer synchronized with the mark on a rotating disc, and it was administered in three sections intended to be of progressive basic difficulty. In the first, trials always proceeded in the same direction in which they started; in the second, sudden changes of direction took place during trials; in the third, there were both sudden changes and momentary stops to arouse the false expectation that changes were occurring.
3. On a sample of 161 prospective air-pilots, the following three measures were obtained: Duration of activity in keeping the ball balanced (Total Time); period during which the pointer was off the mark (Error Time);

duration of *successful* activity (Correct Time, i.e. Total Time minus Error Time). The present study is mainly concerned with the measures " Correct Time " and " Error Time."

4. In Section B subjects tended to make more error on the pursuit-unit and keep the ball balanced for a shorter time than in A. In Section C they made more error, but kept the ball balanced for a *longer* time than in B, and if C is basically a more difficult section than B, then this actual improvement on the one measure has probably come about through the cumulative effect of practice.
5. There is no indication that one of the measures " Correct Time " or " Error Time " is more reliable than the other; as far as can be judged by the inter-section correlations both have a useful reliability.
6. Small but significant positive correlations between " Correct Time " and " Error Time " in Sections B and C are in agreement with the influence of some instrumental dependence between these variates; such dependence implying that the longer the duration of the activity, the greater the scope for accumulating error. In Section A it has probably been overrun by stronger influences germane to the novelty of the situation.
7. " Correct Time," or the duration of successful activity in keeping the ball on the mortar-board, shows no significant relationship with the flying-training criterion. But the aggregate of " Error Time," which is a measure of failure in keeping the pointer synchronized with the rotating disc, has a biserial correlation of —0·35, significant at better than the 7 per cent level.
8. There appear to be no significant differences between the three sections of the test, namely Plain Circuits, Simple Alternations, and Complex Alternations, as regards their respective value for the prediction of success or failure in flying training.
9. Roundabout " Time " (Total or Correct) reflecting the duration of the balancing-activity, relates more noticeably to certain other sensory-motor tests than " Error Time," the largest correlation being —0·35 ($P<0{\cdot}0001$) between Roundabout " Correct Time " and " Time " on Two-hand co-ordination (Moede type). In view of its comparatively small correlation with the other tests and the fact that it has some validity for flying training, Roundabout " Error Time " can possibly make a worthwhile contribution to a battery of such sedentary tests.
10. It is strongly recommended that the present testing procedure be modified by having a constant total duration of performance for all subjects instead of ending a trial when the ball drops. This would eliminate the effect of experimental dependence that benefits those subjects who cannot keep going for long and have proportionately less time in which they can accumulate error.
11. The Roundabout apparatus would lend itself to a variety of experimental work; and the simple system of index-rings for a stop-watch (or clock) might be useful in other tests where the observer has to present stimuli manually at short prescribed intervals.

ACKNOWLEDGEMENTS

Mr. J. B. Kirstein took the photographs. Miss C. M. Elder assisted greatly with computational work. I am also indebted to Mr. A. O. H. Roberts for helpful criticism and suggestions on psychometric points.

REFERENCES

1. DE WET, D. R. Handlebars: a self-paced test of two-hand co-ordination, and some results on air-pilot candidates. *J. Nat. Inst. Personnel Res.*, 1961, 8, 199–208.
2. DE WET, D. R. A variable co-ordination test, and its potentiality as a gauge of aptitude for airmanship. *Psychol. Afr.*, 1961, 9, 86–99.
3. KOERTH, W. A pursuit apparatus—eye-hand co-ordination. *Psychol. Monogr.*, 1922, 31, 288–292.
4. MELTON, A. W. *ed. Apparatus tests: Army Air Forces Aviation Psychology Program Research Reports, No. 4.* Washington, U.S. Govt. Printing Office, 1947.
5. MILES, W. R. A pursuit pendulum. *Psychol. Rev.*, 1920, 27, 361–376.

Mr. D. R. de Wet,
National Institute for Personnel Research,
P.O. Box 10319,
Johannesburg,
Republic of South Africa.

Manuscript received: 22nd August 1961.

CAPE & TRANSVAAL PRINTERS LIMITED, JOHANNESBURG

Reprinted from the Journal of the National Institute for Personnel Research.

J. NAT. INST. PERSONNEL RES., 1960, 8, 122—136

AN IMPROVED STEADINESS APPARATUS AND ITS VALIDITY FOR AIR-PILOT SELECTION

D. R. DE WET

(Received 31st July, 1959)

Equipment for testing steadiness of hand can be classified into three broad types on the basis of a patent behavioural distinction.

The first comprises "Non-volitional" tests, applicable when the main purpose is to study involuntary hand-tremor, usually of toxic, organic, or functional origin. Two well-known examples are Sommer's "Tridimensional Analyser" which uses a system of levers operating three styli to record vertical, longitudinal, and lateral finger-movements; and the Beall and Hall "Ataxiagraph" using an optical system to obtain photographic records of finger-movements in two planes.

In the second group, which can be described as "Volitional-dexterity", the emphasis is on voluntary manipulative performance and special provision is made to eliminate or damp down the effect of finer involuntary tremor. Such tasks are essentially visual-motor co-ordination activities involving the use of one limb. Most tracing-board tests of the kind developed by G. M. Whipple and others come into this category. Throughout the task, the stylus-point is in contact with a solid ground and involuntary movements are greatly reduced.

Between these two extremes of "non-volitional" and "volitional" steadiness there is a wide variety of possible situations compounded from both in different proportions. A good example of such an "Intermediate" test is the standard laboratory apparatus consisting of a stylus and a sloping metal plate having a series of holes graded in diameter. The subject must try to insert the stylus-point almost its full length into each hole and withdraw it without contacting the sides; any contact is recorded as error. The acts of insertion and withdrawal are voluntary, but as the subject works "free-hand" and no part of the stylus is on any solid support, involuntary tremors have unrestricted play and in the smaller holes may become the principal source of error.

Steadiness tests like those used during the last war by the American Army Air Forces (Melton, 1947) for air-crew selection, could also be grouped as "Intermediate". They took the general form of a metal rod supported by the subject at arm's length inside an orifice, as far as possible without contacting the sides. Periods of holding were alternated with periods of rest. The act of holding was voluntary, and here, too, the lack of support meant that involuntary movements were untrammelled, except to a slight extent by the weight of the rod.

Excluding toxic and organic tremors which are usually associated with abnormal physiological conditions, the commonest manifestation of involuntary tremor is that of functional origin. While this may be a recognised symptom of certain psychotic or neurotic conditions, its most familiar occurrence in normal life is in relation to emotional states such as fear and rage.

Steadiness tests of the "Non-volitional" type would seem to be best suited to clinical work and as indicators of temperament traits rather than skill in normal subjects. For occupational selection purposes tests have tended to be of either the "Volitional-dexterity" or the "Intermediate" type in which the degree of latitude allowed for involuntary tremors has been determined to some extent by the specific job requirements and the relative importance attached to the "temperamental" and "skill" aspects.

Let two analogies be compared; say, the respective selection of rifle- and pistol-shooters. The first group will be using a heavier weapon less responsive to fine tremors and held with both hands against other portions of the body to secure a further steadying effect. The latter group will have to hold in one hand and at arm's length a comparatively light weapon with less inertia to dampen unwanted movement and no auxiliary support whatever; a situation of extreme mechanical instability, that not only fosters involuntary tremors but magnifies these greatly at the gun-muzzle. As Melton (1947) records, in connection with findings on an Aiming Stress

1 IMPROVED STEADINESS APPARATUS 1

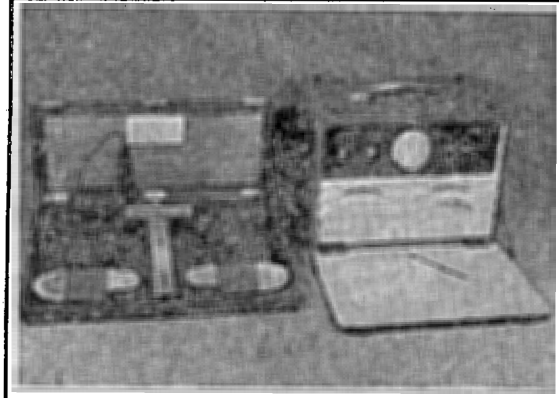


FIGURE A
Subject's Panel and Control Cabinet open for use.

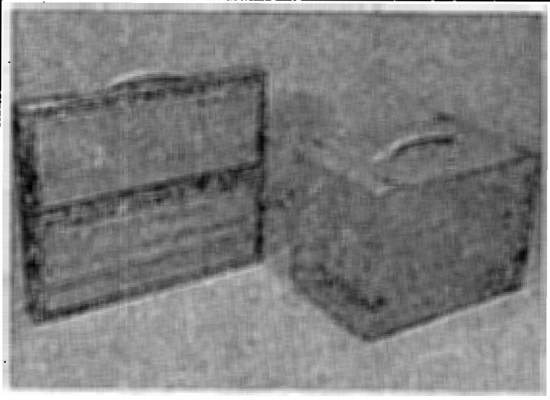
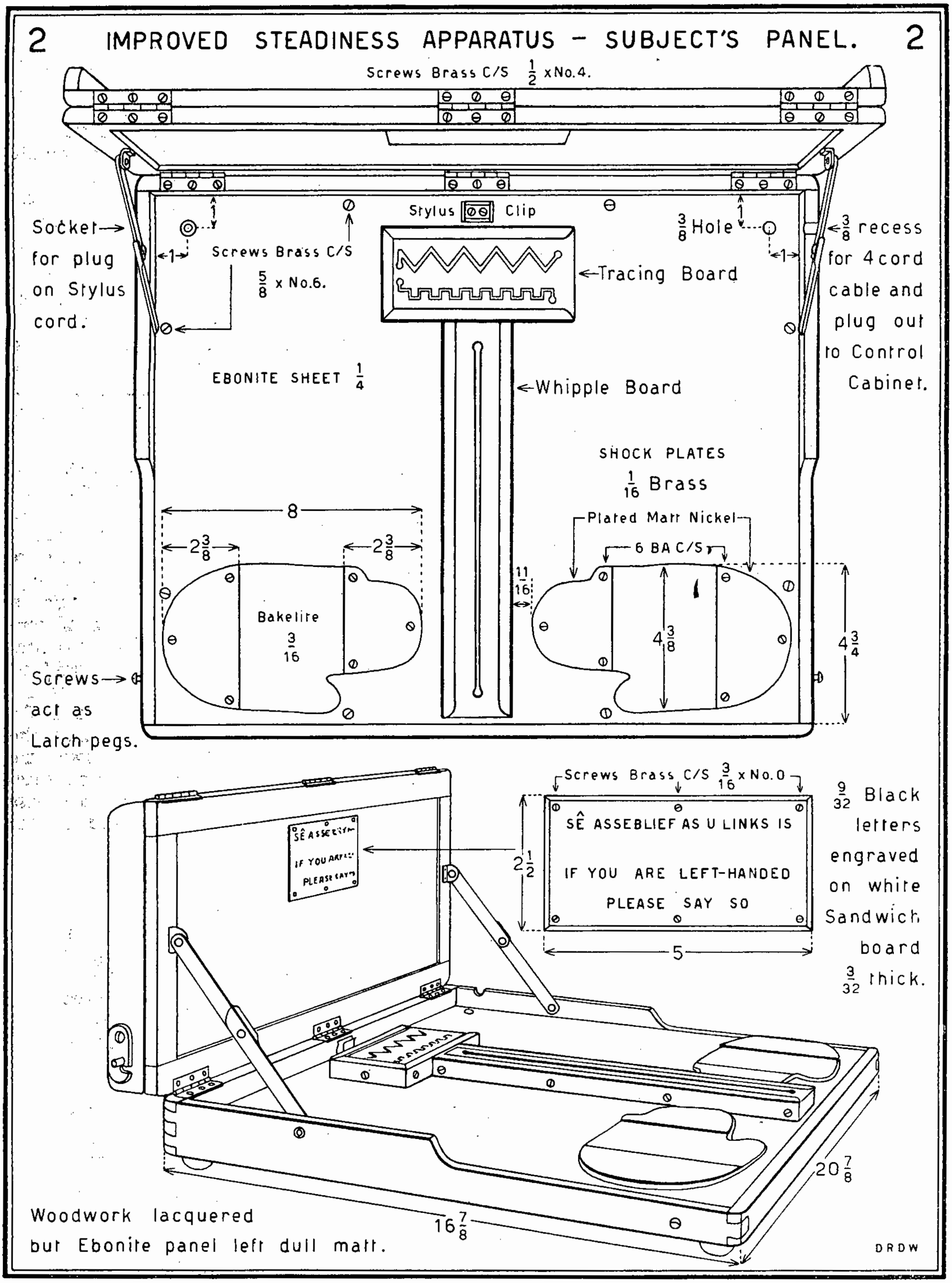


FIGURE B
Units disconnected and closed for transportation.

FIGURE C
Subject doing a 'normal' trial on the Whipple-board. (No threat of shock.)
NOTE: Cord does not usually lie across track.

FIGURE D
Subject doing a 'stress' trial on the Tracing-board. (Hand on shock-plate.)


2 IMPROVED STEADINESS APPARATUS – SUBJECT'S PANEL. 2
Screws Brass C/S 1/2 x No.4.
Stylus Clip
Socket for plug on Stylus cord.
Screws Brass C/S 5/8 x No.6.
3/8 Hole
3/8 recess for 4 cord cable and plug out to Control Cabinet.
Tracing Board
EBONITE SHEET 1/4
Whipple Board
SHOCK PLATES 1/16 Brass
Plated Matt Nickel
6 BA C/S
8
2 3/8
2 3/8
Bakelite 3/16
11/16
4 3/8
4 3/4
Screws act as Latch pegs.
Screws Brass C/S 3/16 x No.0
SÊ ASSEBLIEF AS U LINKS IS
IF YOU ARE LEFT-HANDED PLEASE SAY SO
2 1/2
5
9/32 Black letters engraved on white Sandwich board 3/32 thick.
20 7/8
16 7/8
Woodwork lacquered but Ebonite panel left dull matt.
DRDW

test applied to air-crew candidates: ". . . the arm acts as a lever in the amplification of movement of the upper arm and shoulder musculature".

Moreover, those emotionality factors that relate to a complex fear-pattern and are collectively known to shooters by the sobriquet "Buck-fever" (Libasci, 1949), would be expected to exert greater influence on the performance of the latter group through the superior efficiency of the tremor-transmission.

Hence, while a Steadiness test of the "Volitional-dexterity" type appears to have better face-validity for the rifle-men, something of the "Intermediate" type, with ample scope for involuntary tremor, would seem preferable for the pistol-men.

Irrespective of the different potentialities for recording involuntary movements inherent in the various Steadiness tests, some form of standardised emotion-induction may be included as a portion of the procedure, particularly if selection is intended for occupations in which "stress" situations can arise. In their war-time air-crew selection, the American Army Air Force (Melton, 1947) favoured a "verbal stress". This imposed a distraction-task of having to remember jumbles of numbers and letters simulating combat orders, and was interspersed with stock criticism, admonitions, and running comments delivered very audibly by the observer in the best "back-seat driver" tradition. Under this sort of pressure the subject's emotional set would probably be coloured by rage rather than fear. A consummation by no means inapposite, when it is realised that in addition to the fear-inducing physical hazards of flying, personal relations between instructor and pupil may at times become strained.

The Americans also experimented with powerful auditory distraction including screams and aircraft noises, but did not employ this as a standard procedure.

Prototype Apparatus

In the Aptitude Test Section (A.T.S.) of the South African Air Force, a Steadiness test of the "Volitional-dexterity" type was used. There were two tracing tasks, the first being a standard Whipple-board on which the stylus-point had to be drawn down a straight tapering slot in a metal-plate, and the second a specially designed tracing-board with a zig-zag track imposing changes in direction. The suggestion for supplementing Whipple's straight slot with a track of this latter sort, emanated originally from Miss M. E. White (now Dr. Baehr, University of Chicago) who felt, in common with several others, that the former task by itself might not discriminate adequately between subjects who went through quickly because they had the necessary motor-skill and those who "took a chance" on a quick dash through.

When the subject had performed certain trials "normally", i.e. without any definite emotional pressure, he did a final trial under the threat that he would receive an electric shock if his speed did not at least equal the average standard he had previously attained. Thus on the A.T.S. test, fear rather than rage pervaded the emotional-set.

There were several practical disadvantages in this apparatus. All main components such as boards, switches, shock-plate, induction-coil, counter, and an electronic interrupter providing timed impulses for the latter, were separate units arranged on and below a large table and interconnected with a tangle of flex. The counter recording duration of contacts between the track-sides and stylus-point, and likewise the shocking-system, worked off direct current from accumulators, while the interrupter required alternating current. On both the Whipple- and Tracing-boards the track-plates were firmly screwed down, the former over glass and the latter over bakelite-strip, so that it was a nuisance to replace these backings when they became scored by the stylus-point.

Present Model

The following apparatus retains all behavioural essentials of the A.T.S. prototype, including dimensions of tracks and stylus-diameter, but incorporates extensive technical modifications, mainly with a view to increased compactness, portability, and ease of administration. Apart from the need for reducing the physical inconvenience of the original lay-out, re-design seemed expedient on the psychological rationale that its "make-shift" appearance might adversely affect standardisation by alarming certain subjects and inspiring others with a contemptuous attitude towards the test.

The present model shown in Plate 1, consists of two compact table-units, viz., a Subject's Panel and a Control Cabinet for the observer, interconnected by a four-cord flex. Plates 2 and 3 give the construction of the Subject's Panel and arrangement of the boards and shock-plates. In Plate 3 the lid-stays are described in some detail because commercially-produced fittings of this kind suited to a shallow space are not always easily obtainable. All essential dimensions of the stylus and boards are supplied in Plate 4. The use of a standard No. 53 high-speed drill for the stylus-point facilitates replacement of this item when necessary.

Both tracks are backed with glass-strip held in place by a system of pressure-pads that gives firm holding without the need for screwing down the tracks, and enables the glass to be easily and quickly renewed.

Only about a third of the Control Cabinet shown in Plate 5 is occupied by mechanism and controls; the rest of the space serves to store cable, spare glass, cleaning materials, and testing stationery, etc.

Electrical System

On the original apparatus, a record of the time the stylus has been in contact with the sides of the tracks during a test trial, was obtained by means of an electronic interrupter in conjunction with a reciprocating counter. Later, on a semi-portable demonstration test, some reduction in bulk and slightly better reliability of functioning were obtained by replacing the above electronic device with a mechanical interrupter. This consisted of a serrated disc driven by a small constant-speed motor so that it alternately made and broke contact at regular intervals in the circuit of the counter.

A limitation of both the above systems is that few reciprocating counters will operate satisfactorily at intervals of less than a fifth of a second. On the short tracks in this test, fifths were thought to give rather a coarse measure of performance and tenths became standard, but this entailed a special direct-current counter of considerable size. The present apparatus (Plate 6) records directly on a clock and so dispenses with an interrupter. A suitable model of clock is the "Standard" Type S-6 recording in thousandths of a minute and having its motor wound for 220/230 volts A.C. Its clutch, however, is restricted to 6 volts A.C., because a low voltage ensures less sparking and corrosion between the stylus-point and the track and no risk of harm to any person inadvertently touching such an exposed circuit.

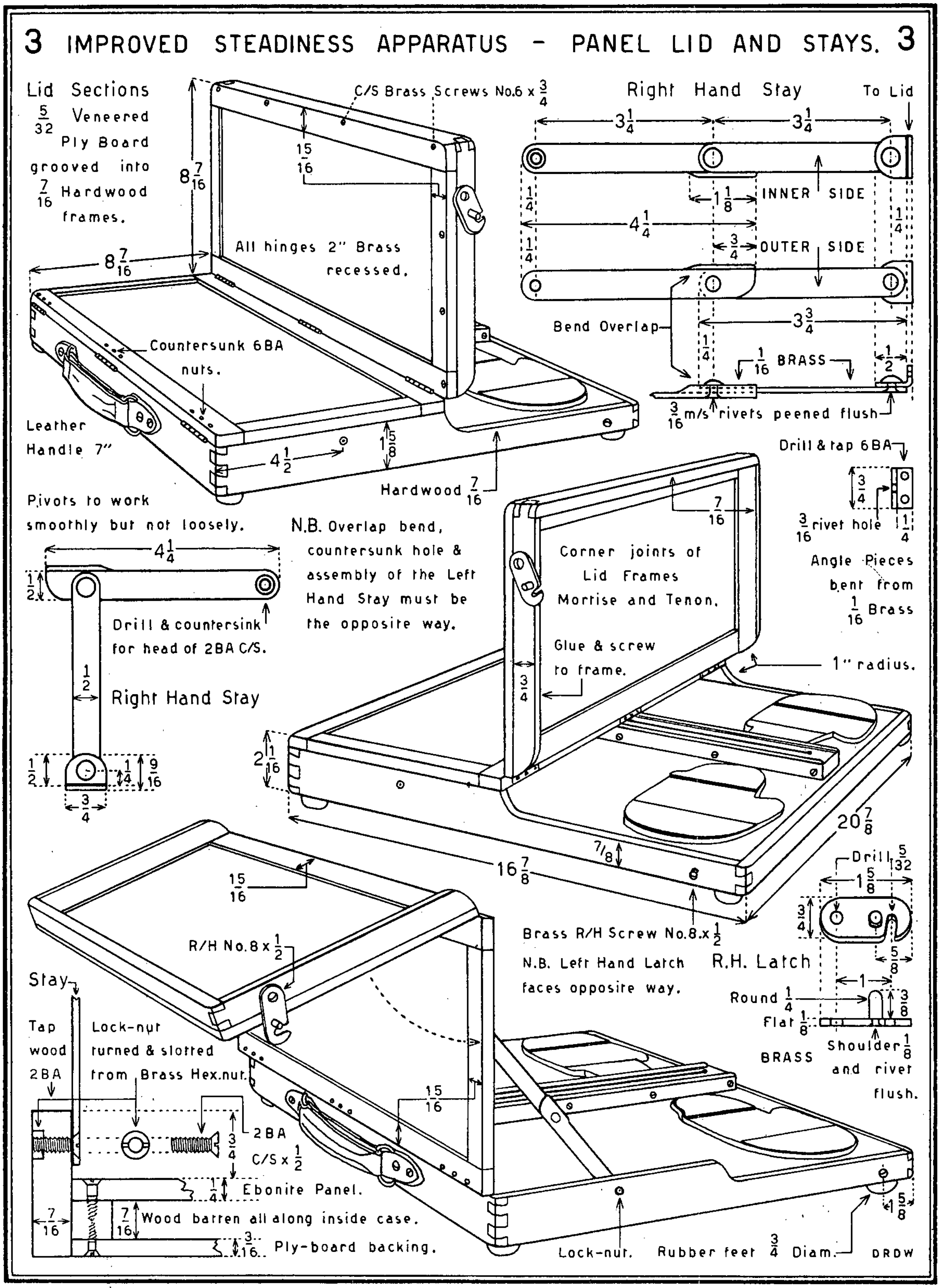
Instead of the separate direct-current induction-coil to supply the shock-potential during the stress trial, there is a simple and compact "built-in" system of two small transformers with their low-voltage sections connected "tail-to-tail" as shown in Plate 6. The high-voltage winding of the one transformer is coupled directly to main current, and the corresponding section on the other becomes a secondary output conveying induced current via a potentiometer and push-button D to the shock-plates. When switch C is closed, a vibrator, which is simply a weak buzzer connected in parallel with the low-voltage poles, comes into operation and gives auditory reinforcement to the threat of possible shock. Actual shock is administered only when push-button D is depressed, and although its intensity may be varied within certain limits by pre-setting the potentiometer, a median setting that gives about 150 volts across the plates is maintained for all subjects. The above method provides a source of induced current high enough in voltage to be decidedly unpleasant to most people, but of sufficiently low amperage to preclude the hazard of injury.

Switch B sets the clock-motor going, but recording can take place only after switch A has been closed. Contact between the stylus-point and track-sides energises the clock-clutch and sounds the buzzer.

Testing Procedure

Figures C and D in Plate 1 show a subject doing the test; and as implied by the small legend inside the lid in Plate 2, he is allowed to use his preferred hand. Holding the stylus like a pencil, and with arm and hand otherwise unsupported, he draws it down the Whipple-track as carefully as possible to avoid touching the sides. Contact records as error, and the longer the duration of any contact, the greater the error. He does this twice, and in addition to his error time, the time taken to complete each run is noted by the observer on a stop-clock.

As a brief change of occupation in lieu of a rest-period, the subject now does eight simple additions on a list of printed numbers, after which he starts on the Tracing-board at the top right-hand corner. Keeping the stylus


3 IMPROVED STEADINESS APPARATUS - PANEL LID AND STAYS. 3
Lid Sections 5/32 Veneered Ply Board grooved into 7/16 Hardwood frames.
C/S Brass Screws No.6 x 3/4
All hinges 2" Brass recessed.
Countersunk 6BA nuts.
Leather Handle 7"
Hardwood 7/16
Right Hand Stay
To Lid
INNER SIDE
OUTER SIDE
Bend Overlap
1/16 BRASS
3/16 m/s rivets peened flush
Drill & tap 6BA
3/16 rivet hole
Angle Pieces bent from 1/16 Brass
Pivots to work smoothly but not loosely.
N.B. Overlap bend, countersunk hole & assembly of the Left Hand Stay must be the opposite way.
Drill & countersink for head of 2BA C/S.
Right Hand Stay
Corner joints of Lid Frames Mortise and Tenon.
Glue & screw to frame.
1" radius.
Drill 5/32
Brass R/H Screw No.8 x 1/2
N.B. Left Hand Latch faces opposite way.
R.H. Latch
Round 1/4
Flat 1/8
BRASS
Shoulder 1/8 and rivet flush.
R/H No.8 x 1/2
Stay
Tap wood 2BA
Lock-nut turned & slotted from Brass Hex.nut.
2BA C/S x 1/2
Ebonite Panel.
Wood batten all along inside case.
Ply-board backing.
Lock-nut.
Rubber feet 3/4 Diam.
DRDW

vertical, and again striving to avoid contact with the sides, he proceeds leftward to the terminal-bay of the top track and then carefully transfers the stylus-point to the adjacent bay of the lower track and traces along to the end on the right. This is followed by another series of eight simple additions, while the observer notes down both the "error" time for these two tracks and the "total" time expended on each. The foregoing procedure in toto constitutes one trial. The subject is informed that the first trial is a practice one and that there is no time-limit for doing it or the next two trials; the object is to work carefully and make as few errors as possible.

However, before starting the fourth or "Stress" trial, he is told the average time he took to complete the second and third trials, and instructed that he must now try to finish within this time, otherwise he will receive a fairly severe electric shock. Procedure is the same as before except that his free hand rests on one or other of the shock-plates throughout the trial.

It is usual to warn subjects that any contact with the sides of the terminal-bays also counts as error.

Application

The test formed part of a selection-battery administered to pupil-pilot candidates for the South African Air Force. Each annual group shown in Table I comprises somewhat less than one third of the whole sample of matriculants, who, having been passed as medically fit, were screened first on the battery and finally by a selection-board.

The Suspensions listed refer only to candidates eliminated because of poor flying-ability; this excludes a few who discontinued training because of medical unfitness that developed on the course or for other reasons.

Results and Discussion

An indication of test-reliability is given by the inter-correlations between trials in Tables II and III; those for the measure "Correct

Table I

SUCCESSFUL PUPILS AND THOSE SUSPENDED FROM TRAINING BECAUSE OF POOR FLYING ABILITY

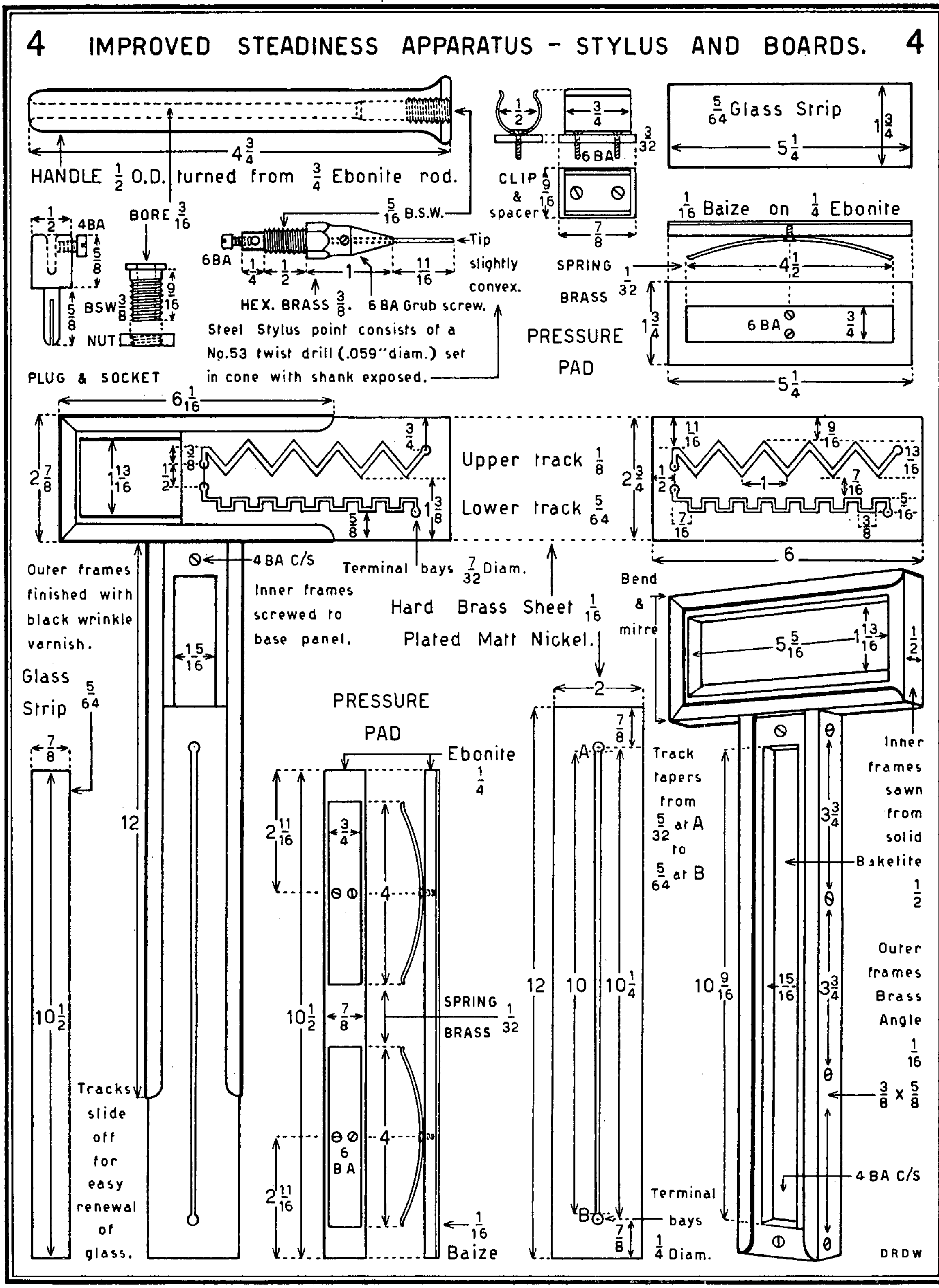
Year	1954	1955	1956	1957	1958
Successes	30	34	33	36	35
Suspensions	8	11	8	11	14
Total N	38	45	41	47	49

The above Total N in each annual group remains the same throughout the five subsequent tables and decimal points are omitted from the correlations and probability-values.

Table II

PRODUCT-MOMENT INTERCORRELATIONS BETWEEN THE FOUR TRIALS ON THE MEASURE "CORRECT TIME" (TOTAL TIME minus ERROR TIME)

TRIALS	1954	1955	1956	1957	1958	MEAN r
1 & 2	77	90	89	93	91	89
1 & 3	73	87	83	89	86	85
1 & 4	56	83	84	86	77	79
2 & 3	95	96	95	97	92	95
2 & 4	76	86	92	89	87	87
3 & 4	83	90	94	87	83	88



4 IMPROVED STEADINESS APPARATUS – STYLUS AND BOARDS. 4
HANDLE 1/2 O.D. turned from 3/4 Ebonite rod.
5/16 B.S.W.
BORE 3/16
PLUG & SOCKET
Steel Stylus point consists of a No.53 twist drill (.059" diam.) set in cone with shank exposed.
HEX. BRASS 3/8. 6 BA Grub screw.
Tip slightly convex.
CLIP & spacer
5/64 Glass Strip
1/16 Baize on 1/4 Ebonite
SPRING BRASS 1/32
PRESSURE PAD
Upper track 1/8
Lower track 5/64
Terminal bays 7/32 Diam.
Outer frames finished with black wrinkle varnish.
4 BA C/S
Inner frames screwed to base panel.
Hard Brass Sheet 1/16 Plated Matt Nickel.
Bend & mitre
Glass Strip 5/64
Tracks slide off for easy renewal of glass.
PRESSURE PAD
Ebonite 1/4
SPRING BRASS 1/32
1/16 Baize
Track tapers from 5/32 at A to 5/64 at B
Terminal bays 1/4 Diam.
Inner frames sawn from solid Bakelite 1/2
Outer frames Brass Angle 1/16
3/8 x 5/8
4 BA C/S
DRDW

Table III

PRODUCT-MOMENT INTERCORRELATIONS BETWEEN THE FOUR TRIALS ON THE MEASURE "ERROR TIME"

TRIALS	1954	1955	1956	1957	1958	MEAN r
1 & 2	80	51	74	68	52	65
1 & 3	64	67	57	56	50	59
1 & 4	68	57	28	39	40	47
2 & 3	69	68	73	87	55	72
2 & 4	67	70	50	56	32	56
3 & 4	50	72	70	68	65	66

Time" (i.e. "Total Time" minus "Error Time") in Table II being appreciably larger than those for "Error Time" (Table III). The inferior reliability of this latter measure could be due to the fact that for a given run "Error Time" is almost invariably a very much smaller proportion of the total than "Correct Time", so that even minor mechanical irregularities like dust or tarnish on the stylus-point and track-sides or scratches on the glass backing-strips can produce substantial variations in the sensitivity of the recording and the standard difficulty of the task.

For this reason observers are urged to change the glass-strips before scratches become too pronounced and to keep the tracks free of dust, occasionally also cleaning their inner edges and the stylus-point with a little fine metal-polish.

In both Tables II and III the mean inter-trial correlations obtained by z-transformations show that the largest of these is between Trials 2 and 3 and the lowest between Trials 1 and 4; which is interesting in view of the fact that correlations between "Error Time" and "Correct Time" in Table IV are generally small.

In Tables II and III, if only the first three "normal" trials are considered, Stress Trial 4 being excluded because of its different motivation, then the smallest mean correlation is between Trials 1 and 3 and the next between Trials 1 and 2. If Trial 1 is excluded and only Trials 2, 3 and 4 are considered, then mean correlations involving Trial 4 remain smaller than the comparable ones for Trials 2 and 3.

The evidence above suggests that Trial 1 and Trial 4 each have some distinctive components additional to the matter they share in common with each other and Trials 2 and 3. Stress Trial 4 is obviously loaded with a fear-motivational influence. Trial 1 can be expected to have elements relating to novelty and adjustability in a strange situation. Some of these may have been tempered by telling the subject that it was a "practice" trial and implying that within reasonable limits he was free to explore for a technique that suited him best in attaining the prescribed goal. Conversely this assurance may have accentuated the effect of other elements by allowing the subject to be more "himself", so that his behaviour in this first trial was very much a reflection of individualism.

In Table IV, "Error Time" was correlated with "Correct Time" for the reason that "Error Time" itself is a part of "Total Time", and a direct correlation with the latter would be biassed by the common element.

The mean correlation of .05 for Trial 1 is not significantly different from zero, so that there seems to be no relation between speed of work and error. Negativity is just apparent in Trials 2 and 3 and becomes significant in Trial 4.

The significant negativity in Trial 4 suggests some tendency for speed of work in this trial to be positively associated with error. A lack of such relationship in Trial 1 may be partly attributable to the collective neutralisation effect of various influences that are not identifiable with the basic motor-skill, but have been engendered by the novelty of the situation. Among these could be counted tension, insecurity, cognitional inertia, social embarrassment, etc., that for a time override whatever

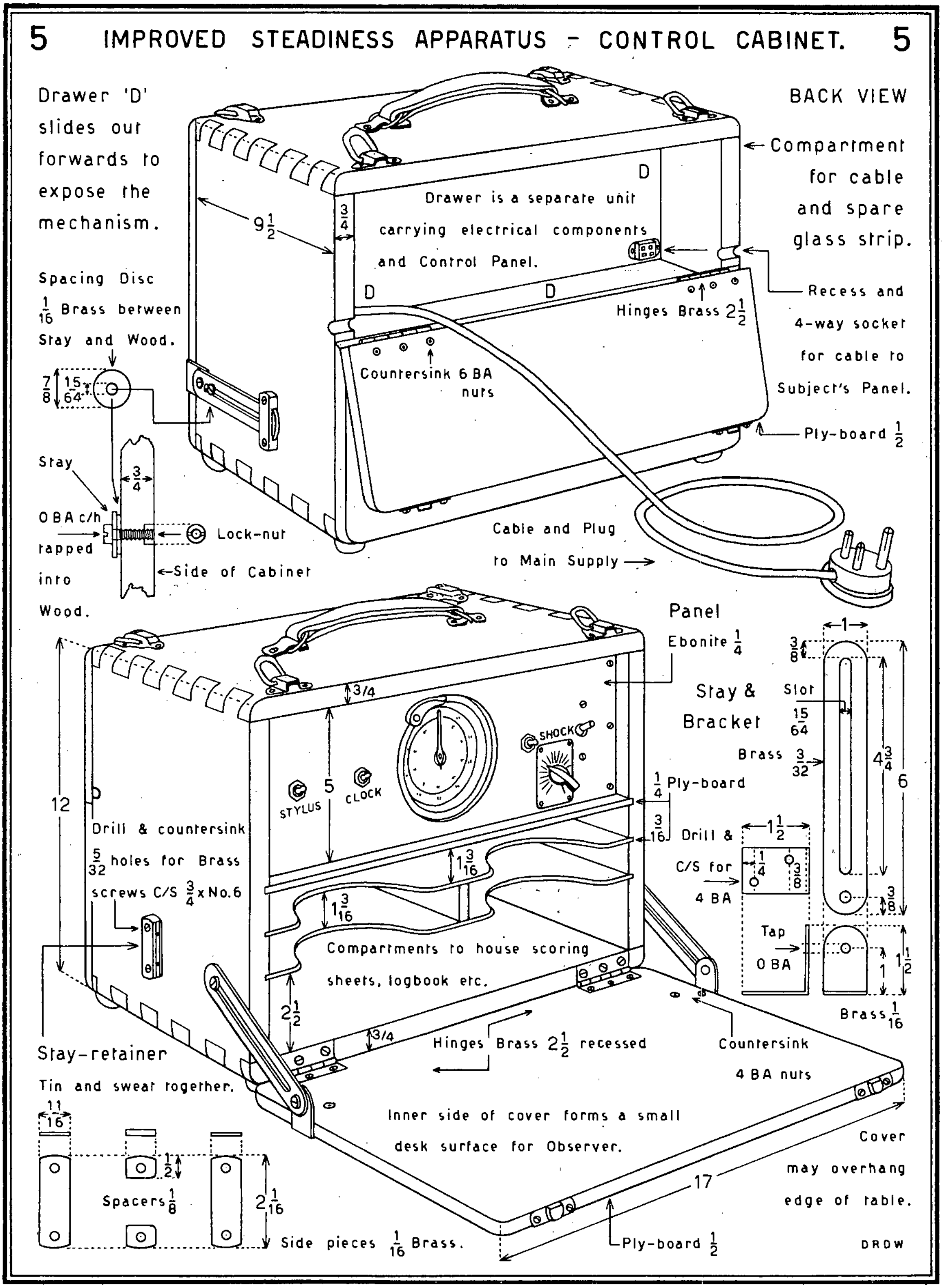

5 IMPROVED STEADINESS APPARATUS - CONTROL CABINET. 5
Drawer 'D' slides out forwards to expose the mechanism.
BACK VIEW
Compartment for cable and spare glass strip.
Drawer is a separate unit carrying electrical components and Control Panel.
D
9 1/2
3/4
Spacing Disc 1/16 Brass between Stay and Wood.
Hinges Brass 2 1/2
Recess and 4-way socket for cable to Subject's Panel.
Countersink 6 BA nuts
7/8
15/64
Ply-board 1/2
Stay
O BA c/h tapped into Wood.
Lock-nut
Side of Cabinet
Cable and Plug to Main Supply
Panel Ebonite 1/4
Stay & Bracket
Slot 15/64
Brass 3/32
4 3/4
6
1
3/8
12
STYLUS
CLOCK
SHOCK
5
1/4 Ply-board
3/16
Drill & C/S for 4 BA
1 1/2
Drill & countersink 5/32 holes for Brass screws C/S 3/4 x No.6
1 3/16
Compartments to house scoring sheets, logbook etc.
Tap O BA
2 1/2
Brass 1/16
Hinges Brass 2 1/2 recessed
Countersink 4 BA nuts
Stay-retainer
Tin and sweat together.
11/16
Inner side of cover forms a small desk surface for Observer.
Cover may overhang edge of table.
Spacers 1/8
2 1/16
17
Side pieces 1/16 Brass.
Ply-board 1/2
DROW

Table IV

PRODUCT-MOMENT CORRELATIONS BETWEEN "ERROR TIME" AND "CORRECT TIME" (TOTAL TIME minus ERROR TIME) FOR EACH TRIAL.

TRIALS	1954	1955	1956	1957	1958	MEAN r	P=
1	23	−18	−14	22	11	05	478
2	06	−20	−36	12	−23	−13	074
3	01	−39	−11	05	02	−09	203
4	−03	−37	−43	−03	−37	−25	0002

(Negative signs imply that fast working on the test tends to be associated with more error.)

potential skill for this task certain subjects may have. More specific to the motor-skill, there is probably a composite abulia, taking the form of varying inability to come to a decision about such aspects as optimum body-posture, muscular equilibrium, angle of arm, strength of grip, downward pressure on the stylus, and rate of progression. Such an influence seems a likely factor behind the overt restlessness shown by some subjects and their proclivity for asking "stupid" questions. It may be aggravated by confused notions about the test which have been gleaned from discussion with others already tested and whose reports are often very much coloured by personal opinion.

When Trial 4 is reached, subjects have settled down to the task, and much of this differential mediation has been reduced; the sensory-motor functions are more stable, and with most of the "uncertainty" resolved, each subject has now a fairly definite and well-established technique. Even more significant is the introduction in Trial 4 of strong, common motivation, represented by the threat of physical punishment if there is failure to attain a certain previous standard.

In Trial 1 there was the desire to excel, diluted with a variety of adventitious influences of which the combined effect was to render the speed-accuracy relationship an individual matter for each subject. In Trial 4, however, these influences are much less, and the initial desire to excel is now braced by common experience and the strong, common motive of wishing to escape punishment. Thus the group becomes more homogeneous in its reaction, as evidenced by the now significant negative correlation between time and error.

A comment by Mr. A. O. H. Roberts puts the matter in a nutshell: "The stress situation generates urgency, and urgency generates error."

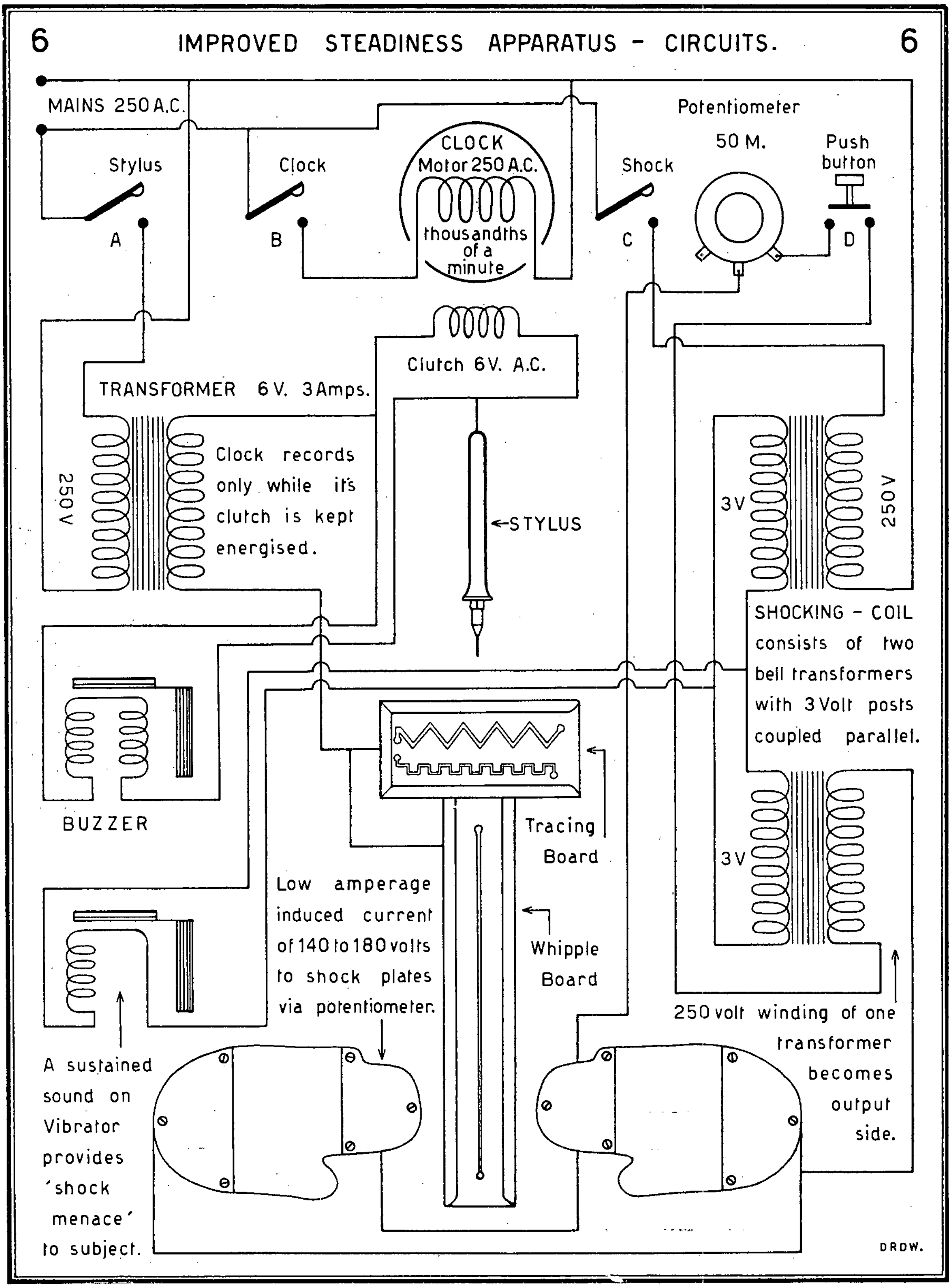
With one exception, the biserial correlations between "Correct Time" and the Success/Suspension flying criterion in Table V are all

Table V

BISERIAL CORRELATIONS BETWEEN "CORRECT TIME" (TOTAL TIME minus ERROR TIME) AND SUCCESS/SUSPENSION IN FLYING TRAINING.

TRIALS	1954		1955		1956		1957		1958		MEAN	P=
	r bis	S.E.	r bis	S.E.	r bis	S.E.	r bis	S.E.	r bis	S.E.	r bis	
1	42	20	32	19	29	21	06	20	−10	20	19	021
2	19	22	25	19	12	22	02	20	07	19	12	124
3	20	22	30	19	05	23	05	20	06	19	12	124
4	30	21	44	17	08	23	15	20	15	19	22	007

(Positive signs imply that success in flying tends to be associated with slow working on the test).


6 IMPROVED STEADINESS APPARATUS - CIRCUITS. 6
MAINS 250 A.C.
Stylus
A
Clock
B
CLOCK
Motor 250 A.C.
thousandths of a minute
Shock
C
Potentiometer
50 M.
Push button
D
Clutch 6V. A.C.
TRANSFORMER 6 V. 3 Amps.
250 V
Clock records only while it's clutch is kept energised.
←STYLUS
3V
250 V
SHOCKING - COIL consists of two bell transformers with 3 Volt posts coupled parallel.
BUZZER
Tracing Board
3V
Low amperage induced current of 140 to 180 volts to shock plates via potentiometer.
Whipple Board
250 volt winding of one transformer becomes output side.
A sustained sound on Vibrator provides 'shock menace' to subject.
DROW.

positive, indicating a tendency for success in flying to be associated with slow working on the test. This appears reasonable, for fast work does not necessarily result from superior motor-skill, but may well be induced by temperamental factors like tension and fluster.

Most of the correlations are small, particularly for the 1957 and 1958 groups. Tate's z*-transformation method was used to compute the mean for each trial over the five annual groups and its level of significance. Trials 1 and 4 reach the 2% level of significance while the other two miss the 10% level. The results of Trials 1 and 4 give some support to the hypothesis that the individual who can remain controlled and unhurried in a strange situation, and likewise in a situation that has become physically threatening, will stand a better chance of succeeding on a course of flying-training than one who lacks these attributes.

In Table VI, the biserial correlations between "Error Time" and the criterion do not differ significantly from zero, which suggests that the qualitative performance on the test bears no relationship to success or failure in flying-training.

as a function of test validity, it has in fact shown less diagnostic value for success in flying-training than the basically similar prototype did during the war. Biesheuvel (1943) used an age-group of below 21.4 years which was thus similar to that in the present samples, and he obtained an average validity of .30.

There is a logical view, supported by Biesheuvel (1958), that the apparent reduction in validity of formerly successful war-time tests when applied to air-pilot candidates during peace-time can to some extent be explained by differences in training procedure. During war the training has of necessity to be highly intensive, impersonal, and standardised very much on a "conveyor-belt" system. The rigid adherence to prescribed programmes allows little extra time or resource to be expended on stragglers who in any way fall short of minimum requirements. On the other hand, in peace-time, instructors have much more opportunity to give personal and specialised tuition; they take extra trouble with pupils' individual shortcomings and so contrive to make successes of some who under war-time conditions would have been suspended without more ado. As a result of this the group accepted for training is made artificially homogeneous and the possible existence of discrimination by test performance becomes obfuscated.

Table VI

BISERIAL CORRELATIONS BETWEEN "ERROR TIME" AND SUCCESS/SUSPENSION IN FLYING TRAINING

TRIALS	1954	1955	1956	1957	1958
1	05	−01	02	03	−03
2	03	04	07	06	−03
3	−03	01	01	03	−12
4	01	−01	−02	06	−12

(Negative signs imply that success in flying tends to be associated with less error on the test). The S.E. of the largest (.12) is .19, hence none differs significantly from zero.

As reported by Melton (1947), the United States Army Air Force found very little significant relationship between Steadiness and Aiming tests and success in air-crew training. A noteworthy feature of their tests is that the measures were almost invariably of qualitative performance, time being a constant for all subjects.

While it is reasonable to believe that the present apparatus is technically more reliable than the prototype, and reliability is accepted

Curtailment through selection-procedures is a possible contributary reason for apparently low validity. The Steadiness test was actually part of the battery on which candidates were screened, and both the "steadiness" score and written comments by the observer on other features of behaviour during this test did carry some weight when assessing the candidate's degree of all-round suitability for enrolment as a pupil.

However, it must be pointed out that these assessments served only as recommendations to assist a selection-board that had full powers to accept or reject them when making its final decisions on which candidates should be admitted for training. As there has often been considerable disparity between recommendations founded on pre-selection and final acceptances, the influence of such curtailment was probably less than could have been expected in a hypothetical case having perfect correspondence between the two groups.

To evaluate more precisely the extent to which discrepancies between recommendations and acceptances affect the apparent validity of tests in a selection-battery, will need further investigation beyond the scope of the present study.

Summary and Conclusions

(1) Details are furnished of a "Steadiness" apparatus basically similar to an earlier prototype, but embodying extensive technical redesign with a view to increased compactness, ease of administration, and reliability.

(2) Of the two performance-measures, "Correct Time" showed greater reliability in terms of inter-trial correlations than "Error Time". On both measures the largest mean correlation was between Trials 2 and 3 and the lowest between Trials 1 and 4; there is also other evidence pointing to the distinctive nature of Trials 1 and 4.

(3) The least significant mean correlation between "Error Time" and "Correct Time" was obtained on Trial 1, and the most significant on Trial 4 (Stress). Hence there is some tendency in Trial 4 for speed of work to be positively associated with error. This appears to be due to the common experience and strong common motivation characterising this stress-trial. The lack of such relationship in Trial 1 may be accounted for by the collective neutralisation-effect of various influences like tension, insecurity, social embarrassment etc., engendered by the novelty of the situation.

(4) "Correct Time" showed some correlation with success in flying-training, and "Error Time", none. The positivity of the correlations with "Correct Time" indicates a general tendency for slow work on the test to favour the chance of success in flying-training; slow work presumably implying an absence of "fluster". This relationship is more noticeable on the first and fourth trials, and it seems likely that in these trials the effect is accentuated by extraneous factors like the capacity for adaptation to a new situation and freedom from disturbance under stress.

(5) Curtailment and aspects of peace-time training have been mentioned as possible contributary reasons for the low average validity of the test. Although it has not hitherto been identified, some progressive change on the training and criterion side during latter years seems not improbable, for such validity as the test showed for the two earlier groups deteriorated markedly during the three later ones; whereas reliability, as indicated by inter-trial correlations, has shown no concomitant reduction. A similar trend has also been observed on certain other tests, and there is no reason to assume that differences in sampling could be entirely responsible.

(6) It has been customary to regard the first trial on some sensory-motor tests as a mere practice or "warming-up" period that is not intended as a measure of serious performances but fulfils the function of setting the subject at his ease, and is thereby supposed to make the later trials more reliable indicators of his actual ability. This derives from perfectly logical considerations to limit variables when the main purpose is to obtain a straightforward measure of sensory-motor ability specific to a particular test and as far as possible uncontaminated with other factors (emotional or otherwise) induced by the novelty of the situation. However, the likelihood should not be ignored that the first trial has particular usefulness in aiding selection for certain occupations, mainly because it is loaded with these very influences.

(7) It is concluded that the sensory-motor ability measured by this test has no appreciable relationship to success or failure in flying-training, and that such small validity as the test does show is attributable rather to the subject's "manner" of application, exemplified by his optimum rate of work, which has probably more affinity with functions of temperament and personality than with basic motor-skill.

(8) When a source of alternating current is available, the present shock-system employing

two small bell-transformers "tail-to-tail" and a suitable potentiometer could also be applied in other laboratory work as a substitute for the more costly and cumbersome induction-coil.

Acknowledgments

Mr. J. B. Kirstein took the two photographs of the test being administered.

Many routine calculations were kindly and competently done by Miss C. M. Elder.

Mr. A. O. H. Roberts was very helpful as a source of expert information on statistical theory and practice. I am also indebted to him for stimulating discussion that contributed towards clarifying the interpretation of several points in the data.

REFERENCES

BIESHEUVEL, S. (1943): The diagnostic value of flying-aptitude tests designed for the South African Air Force. Unpublished report.

BIESHEUVEL, S. (1958): Validity changes in tests with lapse of time and their causes. Paper read at the meeting of the Commonwealth Advisory Council on Defence Science. Melbourne.

LIBASCI, A. M. (c.1949): Buck-fever and the pistol-shooter. Reprinted from The American Rifleman in Askins, C. (1953); *The Pistol Shooter's Book,* Harrisburg, The Stackpole Co.

MELTON, A. W. (1947): Apparatus tests. Report No. 4. Army Air Forces Aviation Psychology Program Research Reports.

TATE, R. F.: The biserial and point correlation coefficients. Univ. N. Carolina Inst. Math. Stat. Mimeographed Series No. 14.

WALKER, H. P. and LEV, J. (1953); *Statistical Inference,* New York, Henry Holt and Co. (Refers to Tate above 269-270).

GOLDFIELDS

IMPROVED STEADINESS APPARATUS

(Note on earlier model)

The internal consistence of performance on the basically similar apparatus used during the war, shown by split-half correlation between errors in the second and third trials on a sample of 342, was 0.677, which, corrected by the Spearman-Brown formula, is 0.807.

Validity, in terms of biserial correlations between a test score of Average total error and success or failure at Flying School, was as follows:

r bis	S.E.	N
0.309	0.067	262
0.361	(Not available)	170

The time measure showed no significant relationship with flying success or failure.

A score consisting of the Average Time divided by the Average Error was investigated as a possible objective temperament measure, the rationale being that the Primary-functioning individual would tend to dash quickly through the task and produce a small score for Av. Time / Av. Error, whereas the Secondary-functioning individual would tend to work comparatively slowly, make few errors, and produce a large score for Av. Time / Av. Error.

On a sample of 67 subjects, who had been rated as Primary or Secondary in accordance with the combined ratings of several testers and the material from two separate questionnaires and an information sheet, a biserial correlation of 0.45 (± 0.06) was obtained between Av. Time / Av. Error and the function assessments. This finding supported the hypothesis. In peace-time, however, the primary-secondary distinction as related to fighter or bomber suitability, was not regarded as very important; the emphasis being more on general pilot suitability.

As regards the apparent lack of validity of the Steadiness Test for post-war selection, Biesheuvel (1) comments thus: "It is not unlikely that the validity of the Steadiness Test during the war was due to the fact that it tapped personality attributes which were of greater importance as determinants of success in flying training at that time than at present."

REFERENCE

(1) BIESHEUVEL, S. Validity changes in tests with lapse of time and their causes. J. Nat. Inst. Personnel Res., 1961, 8 157-169.

Reprinted from the Journal of the National Institute for Personnel Research.

J. NAT. INST. PERSONNEL RES., 1960, 8, 106-116

A PORTABLE HAND-FOOT REACTION TEST

D. R. DE WET

(Received December 12th, 1958)

As an instrument for measuring mere rapidity of voluntary motor-response to sensory stimulation, the Simple reaction-time apparatus is still standard and the most efficient.

The presentation of a stimulus, usually visual or auditory, starts a chronometric recorder which stops as soon as the subject performs the required reaction, and allowing for the small time lag inherent in all electro-mechanical chronoscopes, a fairly accurate measure of reaction time is thus obtainable. In connection with certain modern electronic recorders, it is claimed that the inertia element has been reduced to such an infinitesimal amount that practically absolute measures of reaction-time are possible.

The Simple reaction-time procedure has proved of great value in clinical work, but its limitations in specific occupational forecasting were realised long ago by many of those who pioneered selection for industrial operatives, air-crew, and other personnel.

It very seldom occurs, either in normal routine or emergencies, that situations arise in which a quick simple reaction as an isolated and almost reflex response, can prove efficacious. On the contrary, there may be occasions when such an action precipitates disaster.

Successful reaction in even simple environments depends most often upon some successful antecedent discrimination. This factor of initial cognitive choice increases in importance as activities get more complex, and in an extreme situation like the cockpit of a large aircraft with its multiplicity of controls and instruments, becomes so overwhelming that simple reaction-time is swamped to insignificance. This does not imply that the temporal factor as such no longer carries any weight; it remains very important, but merges into a composite measure where time taken in identifying the correct response from several possible alternatives, is inseparably conjoined to time taken in performing the motor-act of response. As the process of identification or recognition tends to take comparatively much longer than the final efferent response, simple reaction-time becomes a lesser ingredient.

For occupational selection then, far more reliance has been placed on Choice-reaction tests, which vary a good deal in their degree of complexity and nature of stimulus and response, but have as a common feature some *insight* appurtenant to the subject's reactions.

One well-known laboratory apparatus of this type is a direct technical elaboration of the simple-reaction procedure. It takes the general form of three or four response-keys and a small window before the subject, and an equal number of stimulus-keys and a chronoscope behind a screen for the observer. A closed compartment behind the window houses separate but adjacent coloured lamps corresponding in number to the subject's keys. The wiring is so arranged that by depressing a stimulus key the observer causes one of the lamps to come on and the chronoscope to start recording. The subject has to react on the appropriate one of his keys, which extinguishes the lamp and stops the chronoscope. Reaction on any of the other "wrong" keys produces no result and the lamp continues to burn and the chronoscope to record until the correct key is located and depressed.

In an elementary form of this test, which may be called "Immediate Choice-reaction", the subject's response keys are clearly marked in colours corresponding to those of the lights and his main task entails merely direct recognition of the required key to be depressed for whichever colour illuminates the window.

Usually, however, the subject's keys are not marked, and his initial task is one of insight learning, at first with some trial and error, to determine which key relates to a particular light colour.

Chronoscope recordings during this early part of the test are thus mainly measures of learning progress in establishing the association between a colour and a locality. It is only after the subject knows the combination that his performance reflects his Choice-reaction time.

1 PORTABLE HAND-FOOT REACTION TEST. 1

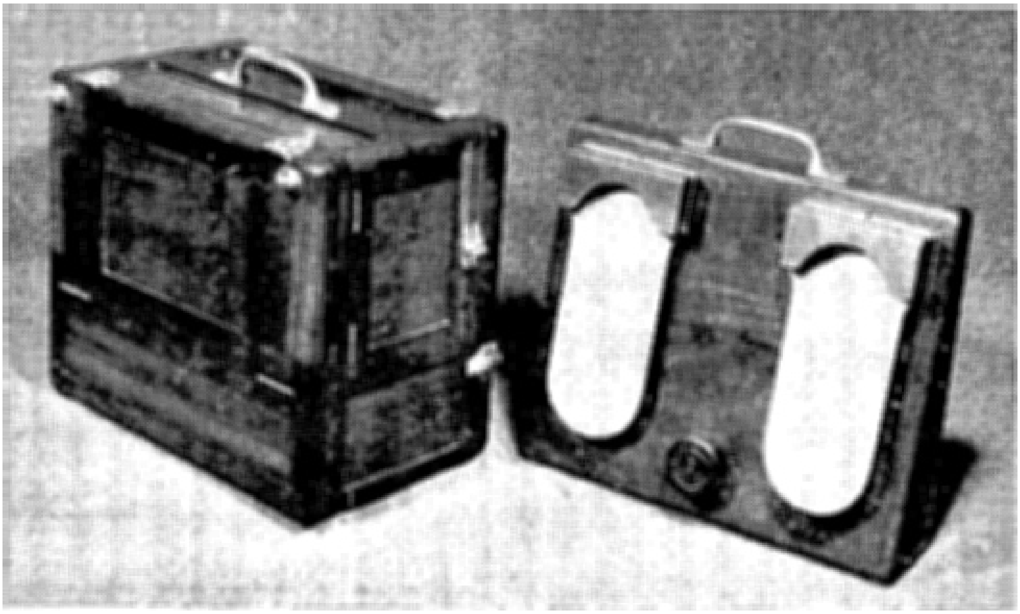


FIGURE 'A' - UNITS DISCONNECTED FOR TRANSPORT.

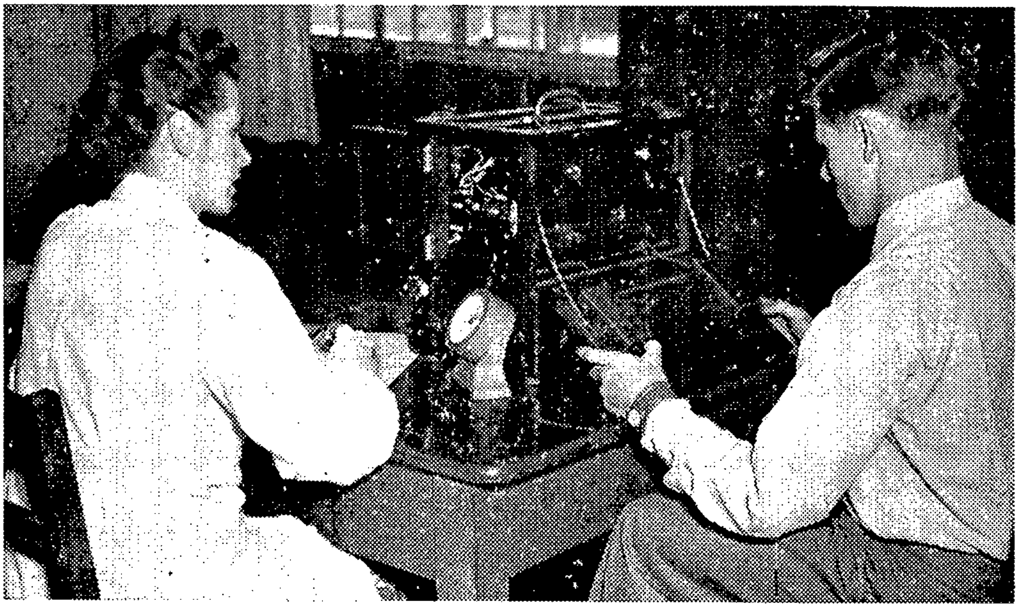


FIGURE 'B' - THE TEST IN OPERATION.

In the above example the stimuli consist of coloured lights that from the subject's view-point are all spatially identical behind the same window. An alternative arrangement is to have separate lamps of the same colour situated in the open at different points on a screen; identification in this case depending not upon distinctive colour but upon physical position.

Moreover, the subject's responses, instead of being restricted to one finger, may be spread out to compass both hands and both feet.

As a technical refinement, the somewhat inconvenient manual presentation of stimuli may be replaced by an automatic system which leaves the observer free to concentrate on studying the subject's general comportment during the test.

For most purposes it appears unnecessarily laborious that discrete reactions be recorded with chronoscopic accuracy. In the learning part of the test, accumulated time covering a series of reactions and measured on an ordinary stop-clock, is quite accurate enough. In the subsequent "speed" part, the number of responses made in each successive period of say half-a-minute is probably quite adequate, provided there is a means of differentiating between correct and wrong responses.

Prototype Models

The apparatus to be described here has no intrinsic originality either as regards fundamentals of Choice-reaction test design or the several variations mentioned above. It comprises no more than an attempt at certain practical improvements on earlier models by other designers.

Behaviourally the present test is identical with a reaction apparatus built in 1941 by Prof. W. Cormack and his assistants in the Department of Electrical Engineering at the University of the Witwatersrand and used for war-time air-crew selection in the Aptitude Test Section of the S.A.A.F. Dr. S. Biesheuvel initiated the construction and advised on the psychological aspects.

This prototype was essentially of a non-portable kind, and as it operated off 24 volts D.C. and suitable rectifiers could not be obtained, a set of bulky accumulators became an indispensable adjunct to the apparatus.

The construction of a semi-portable version of the above was started by P. Webster and completed by G. Walker in the A.T.S. Although less bulky than the original, it also required accumulators.

The condensed outfit depicted in Plate 1 has been designed to fulfil the need for a more mobile Hand-foot reaction test.

General Description

As both military and industrial testing is usually done in urban areas having a main power supply of 220/250 volts A.C., the test was planned *ab initio* to work straight off such current, and thus dispenses with heavy accumulators or rectifiers.

There are two compact, portable units, viz., a table-cabinet and a pedal-board.

The cabinet (Plates 1 & 2) has a hinged front that is lowered to expose four light-jewels on the subject's stimulus-panel and two manual response-keys. At right-angles to this side is a recess housing the observer's control-panel.

The pedal-board (Plate 3) rests on the floor and is connected to the cabinet by a three-cord cable and plug. Another cable and plug leads from the cabinet to the power-supply. For convenience in transport both these cables may be coiled and stored in a compartment at one end of the cabinet (Plate 2).

A larger compartment extending the full length of the cabinet at its base, accommodates supplies of testing stationery.

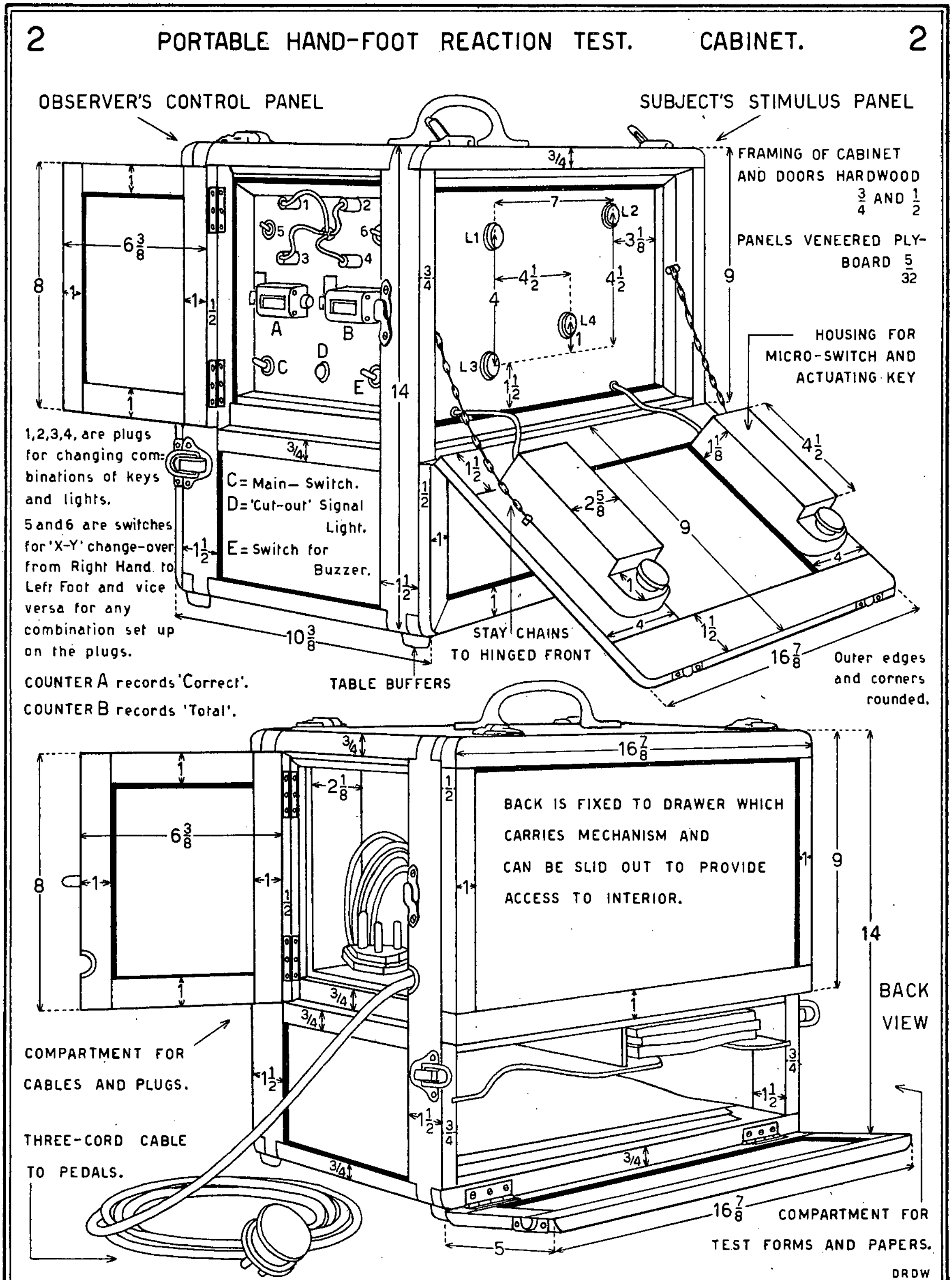
All electro-magnetic components and also the observer's controls are mounted on a separate drawer attached to the back panel and can be slid out as an entire unit for maintenance purposes.

Testing Procedure

The subject is seated comfortably with his hands at the manual-keys and his feet on the pedals. He is then instructed that each one of these four controls relates to a particular light, and if he presses the correct control for the given light, a buzzer sounds and he scores a point.

The light sequence is random, only one comes on at a time, and it remains on until the subject presses and releases any one control (right or wrong), when it is extinguished and the next light-stimulus appears.

His task is to find out by trial and error the control that always makes a given light


2 PORTABLE HAND-FOOT REACTION TEST. CABINET. 2
OBSERVER'S CONTROL PANEL
SUBJECT'S STIMULUS PANEL
FRAMING OF CABINET AND DOORS HARDWOOD $\frac{3}{4}$ AND $\frac{1}{2}$
PANELS VENEERED PLY-BOARD $\frac{5}{32}$
HOUSING FOR MICRO-SWITCH AND ACTUATING KEY
1,2,3,4, are plugs for changing combinations of keys and lights.
5 and 6 are switches for 'X-Y' change-over from Right Hand to Left Foot and vice versa for any combination set up on the plugs.
COUNTER A records 'Correct'.
COUNTER B records 'Total'.
C = Main– Switch.
D = 'Cut-out' Signal Light.
E = Switch for Buzzer.
STAY CHAINS TO HINGED FRONT
TABLE BUFFERS
Outer edges and corners rounded.
BACK IS FIXED TO DRAWER WHICH CARRIES MECHANISM AND CAN BE SLID OUT TO PROVIDE ACCESS TO INTERIOR.
BACK VIEW
COMPARTMENT FOR CABLES AND PLUGS.
THREE-CORD CABLE TO PEDALS.
COMPARTMENT FOR TEST FORMS AND PAPERS.
DRDW

"buzz", and to go on applying it until he can make twenty successive correct responses; which constitutes the criterion of having learnt the combination. He must accomplish this within a time limit of fifteen minutes.

In the "speed" part of the test he is told to continue pressing the correct controls for the lights as he has learnt to do, but this time without the corroboration of the buzzer, which is switched off at E, Plate 2.

The combination remains the same and he must work as quickly and accurately as possible.

After a practice trial of one minute he does a continuous run of five minutes during which the observer records his response-scores on the "Correct" and "Total" counters (Plates 2 & 5) at every half-minute interval.

If the subject depresses more than one control at a time in either part of the test, the apparatus "cuts-out" and the observer has to reset it by putting the main-switch C (Plates 2 & 5), off and on again.

Application

Like its prototypes, the test is being used mostly for pupil-pilot and navigator selection. However, it has also shown promise for other categories, such as wireless operators, research technicians, scientific administration staff, and clinical consultation cases.

On two consecutive groups of pupil-pilot candidates, the following biserial correlations were obtained between performance on the test and the Success/Suspension flying criterion:

	Year	1954	1955
	N	38	47
Learning Time	r_{bis}	.31	.52
	r_{bis}/S.E.	1.45	3.48
Reaction Time	r_{bis}	.60	.48
	r_{bis}/S.E.	3.54	3.13

A rough check in which the above correlations were assumed to be estimates of product moments, showed the significance of the differences between comparable ones for the 1954 and 1955 groups to be well outside the 10% level. It is likely that the differences reflect only sample fluctuations.

In view of the small samples involved, Tate's z*-transformation for testing the significance of biserial differences was not considered to be usefully applicable.

As a possible means of increasing the test's predictive value, a modification was introduced and applied to some later groups of pupil-pilot candidates (not shown here). This had its origin in the assumption that visual stimuli are generally more important than auditory to the aircraft pilot.

A small green light-jewel was mounted on the subject's panel approximately central to the four lights and connected up in place of the buzzer. Correct response to a given light was now indicated by an accompanying illumination of the green jewel instead of a sound on the buzzer.

This change failed in its object, and the author's theory, in so far as an increase of test usefulness was concerned, appears to have been refuted by the results of Van der Reis (1958).

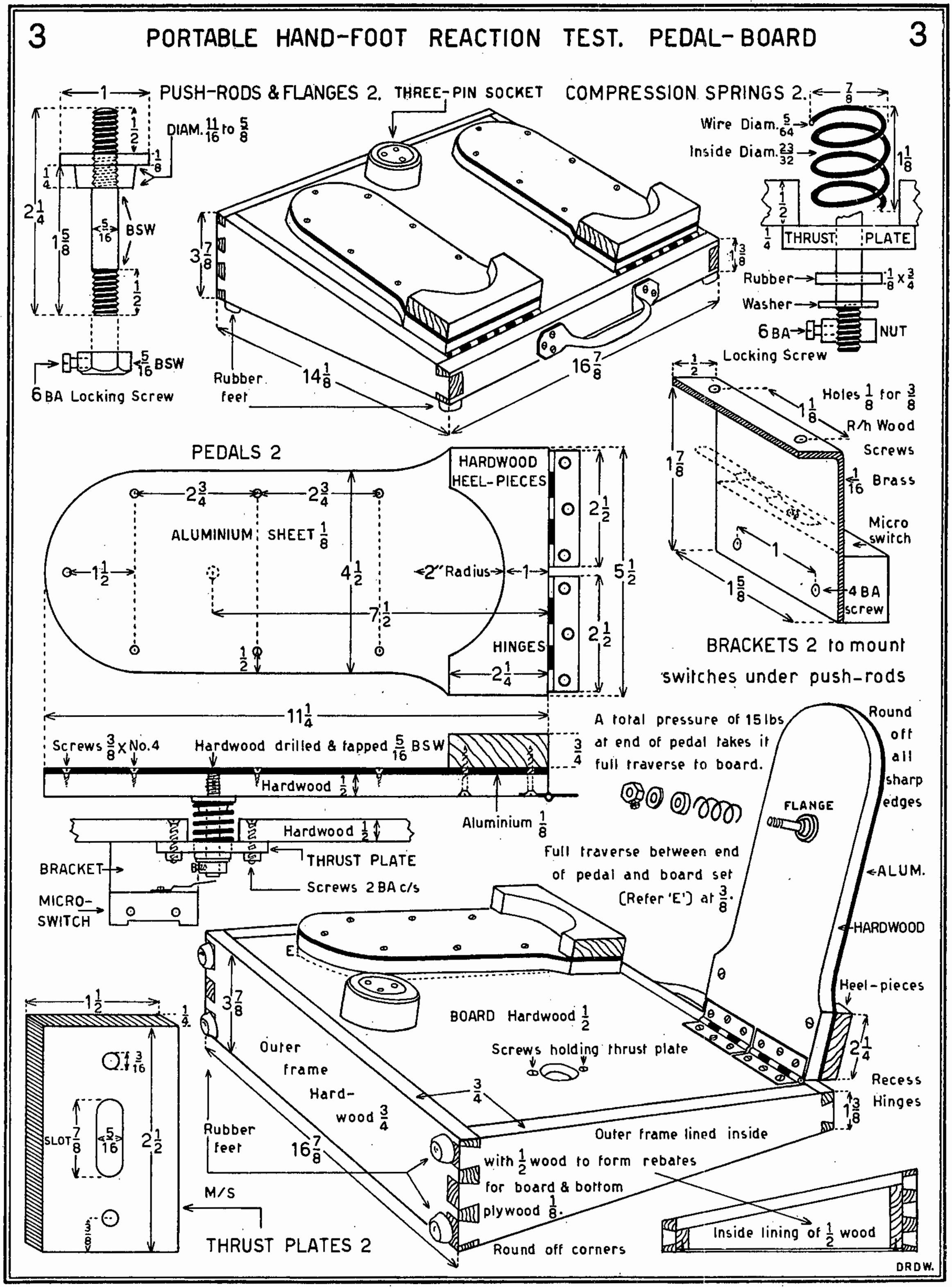
She found that under the new system neither the learning nor the reaction part of the test was significantly correlated with the flying criterion, and moreover there was a far greater preponderance of subjects who were unable to learn the combination within the time limit.

To account for this increase in test difficulty, she advances several astute explanations, of which the most noteworthy refers to additional time taken in eye-movements between stimulus and indicator lights.

However, the possibility of a more straightforward explanation should not be ignored, namely, that the addition of a fifth light to an already complex system of four made the test instructions more difficult for subjects to understand, and a larger number tended to flounder on with an imperfect knowledge of what was really required of them.

Unfortunately the exigencies of testing for selection purposes have as yet precluded any investigation on the various theoretical issues that might relate to this apparent superiority of the auditory indicator.

In any case, loss of validity was considered ample justification for restoring the *status quo;* the green light-jewel was accordingly removed and the buzzer re-coupled as indicator of correct response.


3 PORTABLE HAND-FOOT REACTION TEST. PEDAL-BOARD 3
PUSH-RODS & FLANGES 2.
THREE-PIN SOCKET
COMPRESSION SPRINGS 2.
DIAM. 11/16 to 5/8
BSW
6BA Locking Screw
Wire Diam. 5/64
Inside Diam. 23/32
THRUST PLATE
Rubber 1/8 x 3/4
Washer
6BA NUT
Locking Screw
Rubber feet
Holes 1/8 for 3/8 R/h Wood Screws
1/16 Brass
Micro switch
4BA screw
BRACKETS 2 to mount switches under push-rods
PEDALS 2
HARDWOOD HEEL-PIECES
ALUMINIUM SHEET 1/8
2" Radius
HINGES
Screws 3/8 x No. 4
Hardwood drilled & tapped 5/16 BSW
Hardwood 1/2
Aluminium 1/8
A total pressure of 15 lbs at end of pedal takes it full traverse to board.
Round off all sharp edges
FLANGE
BRACKET
MICRO-SWITCH
THRUST PLATE
Screws 2BA c/s
Full traverse between end of pedal and board set (Refer 'E') at 3/8.
ALUM.
HARDWOOD
Heel-pieces
BOARD Hardwood 1/2
Screws holding thrust plate
Recess Hinges
Outer frame Hard-wood 3/4
Rubber feet
Outer frame lined inside with 1/2 wood to form rebates for board & bottom plywood 1/8.
Round off corners
Inside lining of 1/2 wood
SLOT 7/8
M/S
THRUST PLATES 2
DRDW.

Selector System

Automatic presentation of the stimulus-lights in a predetermined and fixed random sequence is effected by the selector-unit in Plate 4, which constitutes the principal electro-mechanical component.

The earlier models of the apparatus had utilized standard telephonic line-selector switches working off 12 or 24 volts D.C.

As no industrial product of this sort designed for 220/250 volts A.C. could be obtained, it was necessary to construct one.

The solenoids and frame-extensions were cut off a standard line-selector, and the contact-banks and ratchet-mechanism fitted with a laminated armature and a new spring-retainer bent from light steel-bar.

This assembly is supported over a bakelite sub-base by three metal pillars having their top ends threaded and provided with nuts, so as to permit fairly accurate three-point adjustment for height and both lateral and fore-and-aft tilt.

A special magnet with a laminated core made from transformer-iron, is mounted below the armature, and a suitable distance and alignment between the two obtained by adjustments of the pillar-nuts.

Not only is the overall armature-traverse important for satisfactory working of the ratchet-mechanism, but both contact-faces must be accurately "squared on" to prevent excessive vibrations and humming noises which subjects might confuse with the sound of the buzzer.

For compactness, the "cut-out" relay that switches off all stimulus-lights if subjects depress more than one control at a time, is mounted on a small bracket under the pawl of the selector-switch; although there is no direct mechanical coupling between them which makes this location compulsory.

There are two semi-circular contact-banks, each having twenty-five insulated contacts engaged in turn by a brush on the ratchet-spindle.

When the magnet-coil is energised, the armature is drawn down against spring-tension and the pawl raised from one ratchet-tooth on to the next. When the current is switched off, the armature is drawn up by spring-tension and it causes the pawl to drag the ratchet round sufficiently for the two brushes to engage their next pair of contacts.

Electrical Circuits

Referring to Plate 5, the twenty-five contacts on Bank No. 1 are wired in the random order shown here to form four groups, each linking up permanently with one of the stimulus-lights L1, L2, L3 and L4.

Now in exactly the same order, the twenty-five contacts on Bank No. 2 are wired up into four groups identically parallel with those on Bank No. 1, and each of these is connected, preferably by a wire of distinctive colour, to one of the four single-pin plugs 1, 2, 3 and 4.

Each of the four response-controls, viz., the two manual-keys marked Left H and Right H, and the two pedals Left F and Right F, is basically a micro-switch (normally open) in series with the solenoid of a triple-pole relay having at least one pole double-throw. These relays are marked LH, RH, LF and RF respectively, to correspond with their controls.

The "cut-out" relay (mounted for convenience on the selector-unit, Plate 4) is shown at the bottom of Plate 5. It is also triple-pole, but with at least two poles double-throw.

One main 220/230 volt input lead is common to all of the following: The primary coils of the 6 volt "Light" transformer (lower right) and the 6 volt "Buzzer" transformer (middle); the solenoids of the four "control" relays and the "cut-out" relay; the solenoids of the two re-set counters A and B (lower left) and the selector-magnet (top).

The other 220/250 volt input lead passes via the main-switch C to the primary of the "Light" transformer and the first and third "arm" contacts of the "cut-out" relay. From the latter pole, which is normally closed, current is fed through to the four response-controls and the first and second "arm" contacts on each of the four "control" relays LH, RH, LF and RF.

The third sections of these four relays and the second section of the "cut-out" relay are arranged to give series continuity between the one 6 volt secondary pole of the "Light" transformer (lower right) and the selector-brush of Bank No. 1 regulating distribution to the four stimulus-lights L1, L2, L3 and L4.

From the other 6 volt secondary pole of the "Light" transformer a common leads to the four stimulus-lights and the "cut-out" signal-light D on the observer's control-panel.

The second sections of the two "control" relays LH and LF lead respectively to the upper and lower movable contacts of the double-pole-double-throw switch No. 5, and corresponding sections of the "control" relays RH and RF to the movable contacts of a similar switch No. 6.

To accommodate the four single-pin plugs 1, 2, 3 and 4 of Bank No. 2, four sockets are provided. Each of these is connected as shown to one fixed contact on switch No. 5 and one on switch No. 6, so that the following settings are established:

(a) With both switches in the "up" or "X" position, relationship between relays and sockets is: LH to Left Upper; RH to Right Upper; LF to Left Lower; and RF to Right Lower.

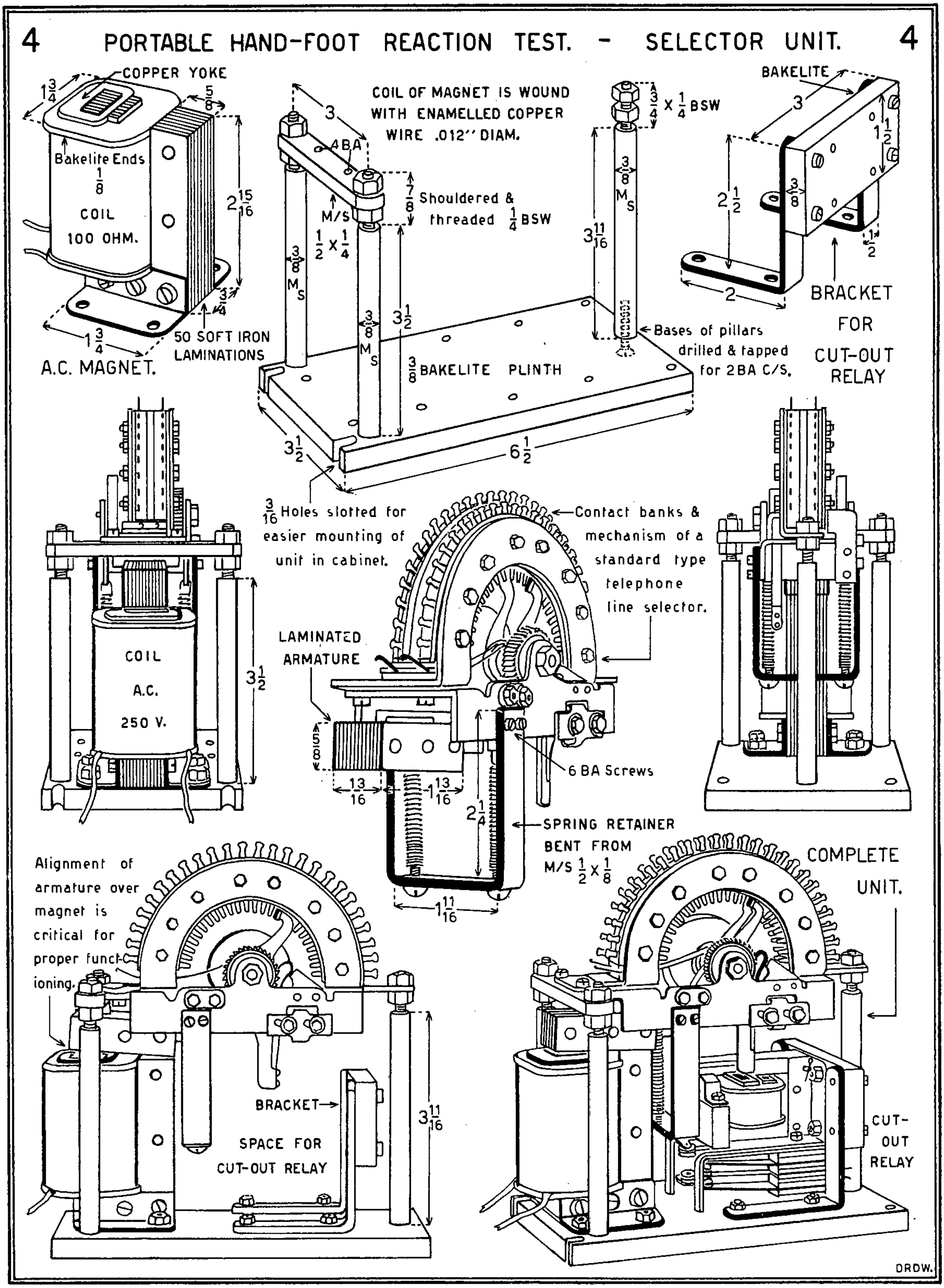
(b) With both switches in the "down" or "Y" position, relationship between relays and sockets is: LH to Right Lower; RH to Left Lower; LF to Right Upper; and RF to Left Upper.

By this means, as each of the four response-controls is coupled permanently to its own relay, the alteration of switch-setting from "X" to "Y" produces a "diagonal" interchange between controls and sockets: Right-hand takes the place of Left-foot; Left-hand that of Right-foot, and vice versa.

As was mentioned earlier, the four plugs are wired to distinct groups of contacts on Bank No. 2 positionally identical to their counterparts on Bank No. 1.

A lead passes from the brush of Bank No. 2 to the "Correct" counter A and one primary pole of the "buzzer" transformer.

Now although the *sequence* of light-presentation remains fixed in accordance with the grouping on


4 PORTABLE HAND-FOOT REACTION TEST. - SELECTOR UNIT. 4
COPPER YOKE
Bakelite Ends 1/8
COIL 100 OHM.
50 SOFT IRON LAMINATIONS
A.C. MAGNET.
COIL OF MAGNET IS WOUND WITH ENAMELLED COPPER WIRE .012'' DIAM.
4BA
M/S
7/8 Shouldered & threaded 1/4 BSW
1/2 x 1/4
3/4 x 1/4 BSW
3/8 BAKELITE PLINTH
Bases of pillars drilled & tapped for 2BA C/S.
BAKELITE
BRACKET FOR CUT-OUT RELAY
3/16 Holes slotted for easier mounting of unit in cabinet.
Contact banks & mechanism of a standard type telephone line selector.
LAMINATED ARMATURE
COIL A.C. 250 V.
6 BA Screws
SPRING RETAINER BENT FROM M/S 1/2 x 1/8
Alignment of armature over magnet is critical for proper functioning.
BRACKET
SPACE FOR CUT-OUT RELAY
COMPLETE UNIT.
CUT-OUT RELAY
DRDW.

the selector-banks, the combination of correct controls and lights may be varied by inserting the plugs into different sockets. The criterion of correct response for any given combination is that the "buzzer" transformer and the "Correct" counter A shall be potentialized via the brush of Bank No. 2. (In the first or learning part of the test, with switch E closed, the buzzer itself is also potentialized and the resultant sound is the subject's criterion of correct response.)

To ensure that the controls and lights are switched off automatically if the subject depresses more than one control simultaneously, a principle conceived by Prof. Cormack and applied on his prototype, is used here with minor modifications to suit the present circuit.

An extension is taken from the coupling between each response-control and its relay-solenoid to a carbon resistor of 9000 ohms. These four equal resistors have their other ends common with one lead to the solenoid of the "cut-out" relay.

When any one response-control is depressed, the resistance of 9000 ohms provides an adequate barrier to the solenoid of the "cut-out" relay; but with two or more controls depressed at the same time, two or more of the resistors in parallel let sufficient current through to trip the "cut-out". When this happens the first section of contacts closes, current from the main by-passes the resistors, and the relay "hangs in". The second section breaks continuity to all stimulus-lights and flashes on the observer's signal-light; the third section breaks the common to all four response-controls.

By turning off main switch C, the "hanging-in" circuit is broken and the "cut-out" relay returns to its normal, unenergised position.

The magnet of the selector-unit as described in Plate 4 was found to be somewhat more powerful than necessary. A wire-wound resistance of about 850 ohms, .5 amp. (not shown) was accordingly fitted in series and made operation more quiet.

Although, for the sake of clarity, the selector-brushes in Plate 5 are shown as consisting of one arm apiece, in practice each has two arms at 180 degrees, so that stepping proceeds directly from the last bank-contact in the series on to the first again.

The following brief synopsis of functions completes the circuit description:

When the main switch C is closed, one of the four stimulus-lights (determined by the contact on which the selector-brush of Bank No. 1 happens to be at rest) comes on.

As the fixed contacts on the first sections of the four "control" relays are in common with the selector-magnet and "Total" counter B, and the third sections are links in series with the brush, actuation of any relay by depressing its appropriate control energises the selector-magnet, records once on the "Total" counter, and extinguishes the stimulus-light.

If the control depressed and the light are correct partners for the combination pre-set on the "X" and "Y" switches and the plugs, the "Correct" counter A also records once, and provided that switch E is closed, the buzzer sounds.

If control and light are not correctly matched by the subject for the given combination, there can be no continuity to the "Correct" counter and the buzzer.

When the control is released, all solenoids mentioned above are de-energised and the selector-mechanism moves on another step to illuminate another stimulus-light; or occasionally, the previous light again.

Should more than one control be depressed at a time, the "cut-out" process described earlier takes place and the main switch C must be turned off and on again to reset the circuits.

Discussion

The four stimulus-lights, like those on the prototype models, are arranged unsymmetrically to prevent a too direct spatial association with the controls.

There are twenty-four possible combinations of lights and controls. The following four of these, which were found in a short experimental study to be pretty nearly equal in difficulty, have proved adequate for general selection testing:

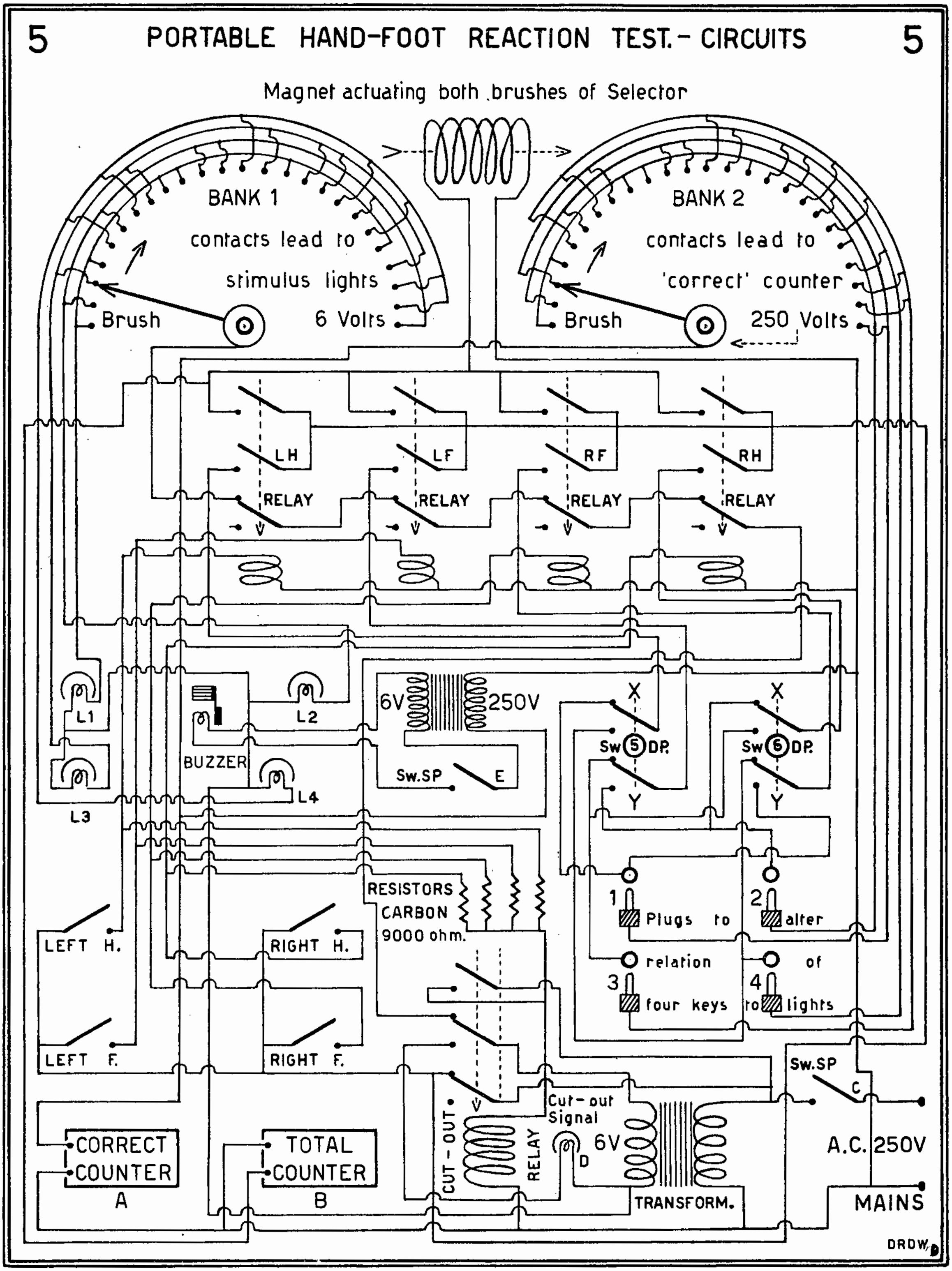
(1)	L1 Right foot.	L2 Left hand.
	L3 Left foot.	L4 Right hand.
(2)	L1 Left foot.	L2 Left hand.
	L3 Right foot.	L4 Right hand.
(3)	L1 Right hand.	L2 Right foot.
	L3 Left foot.	L4 Left hand.
(4)	L1 Left hand.	L2 Left foot.
	L3 Right foot.	L4 Right hand.

A significant behavioural feature of the learning part on this test and its predecessors, and one that distinguishes them from more usual laboratory choice-reaction tests, is that the subject is permitted only a single immediate response-essay to a given light. Whether he is right or wrong the light promptly goes out. To identify a correct response he has to rely on comparatively extraneous auditory indication.

On more usual tests of this kind, the light remains on until the correct control is located by trying more than one in succession, and here the criterion of correct response is the going out of the light.

At present, available data does not warrant any speculation on the relative advantages of the two systems for application to highly selected groups aspiring to complex occupations like aircraft-pilotage.

However, it appears likely that the latter method would be preferable in a test intended for subjects less advanced intellectually or culturally, for there is no doubt that the


5 PORTABLE HAND-FOOT REACTION TEST.- CIRCUITS 5
Magnet actuating both brushes of Selector
BANK 1
contacts lead to
stimulus lights
6 Volts
Brush
BANK 2
contacts lead to
'correct' counter
250 Volts
Brush
LH
RELAY
LF
RELAY
RF
RELAY
RH
RELAY
L1
L2
L3
L4
BUZZER
6V
250V
Sw.SP
E
X
Sw 5 DP.
Y
X
Sw 6 DP.
Y
RESISTORS
CARBON
9000 ohm.
1
2
Plugs to alter
relation of
3
4
four keys to lights
LEFT H.
RIGHT H.
LEFT F.
RIGHT F.
Sw.SP
C
CORRECT COUNTER
A
TOTAL COUNTER
B
CUT-OUT
RELAY
Cut-out Signal
6V
D
TRANSFORM.
A.C. 250V
MAINS
DRDW

former imposes more difficult tasks, both in understanding the nature of the criterion and in learning the combination.

Many less successful subjects seem to be baffled by the fact that the *lights themselves* behave in exactly the same way whether the control depressed is right or wrong. Such an attitude could be induced to a considerable extent by prior conditioning, founded on the very definite relationship in everyday life where lights go out only when their proper switches are turned off. Further research is necessary to determine how far the usefulness of the test depends on the influence of "perseveration" or "rigidity" which affects those subjects who are so hide-bound by well-fixated associations, that they are hampered in acquiring new sets.

It was with this in view that the "X" and "Y" switches were fitted to the apparatus. On the prototype model, there was provision for altering light and control combinations by four binding-posts and terminals, but no "X" and "Y" change-over system.

The inclusion of some such quick change-over that would give a convenient "diagonal reversal" of any combination set up at the distribution-points, was later suggested by Dr. Biesheuvel as a worth-while addition for investigating "perseveration".

Having learnt one combination thoroughly, subjects may be switched over to another diagonally opposite, and some measure obtained of the hindrance-effect exercised by the "X" over the "Y" activity.

Summary and Conclusions

(1) A reaction apparatus retaining the same behavioural aspects found of value in two earlier models, but incorporating certain technical modifications and improvements, is fully described.

(2) It has given reasonably trouble-free service for a period of ten years, during which approximately 2,750 subjects have been tested.

(3) The ideal of compactness has possibly been carried a little too far, and internal components, although accessible, are somewhat more crowded than is desirable for ready inspection and maintenance.

(4) The frame-and-panel design of the cabinet provides a structure both light and strong but one which does not lend itself to easy or economical reproduction in any quantity. A slightly larger and plainer cabinet would be an advantage, despite the increase in weight.

(5) An apparatus reverting to the prototype system, with all electro-magnetic components working off 24 volts D.C. and used in conjunction with one of the modern selenium-rectifiers that are today commercially obtainable, is worth consideration. This will obviate the need for making a special selector-unit, but increase the overall bulk and weight, because in addition to the rectifier, a transformer of substantial size will be required.

Acknowledgments

The biserial correlations quoted are from among those computed by Mrs. A. P. Van der Reis.

Mr. J. B. Kirstein took the photograph of the test being administered to a subject by Mrs. V. R. Wilson.

Mr. D. J. M. Vorster fitted the micro-switches to the four controls of the apparatus, and these have proved far more positive and reliable than former sets of plain nickel contacts.

Mr. A. O. H. Roberts was most helpful in connection with statistical queries.

REFERENCE

VAN DER REIS, A. P. (1958): A technical note on the reaction time test. Internal report, Nat. Inst. Personnel Res.

PORTABLE HAND-FOOT REACTION TEST

(Note on earlier model)

The reliability of the earlier war-time model, which incorporated the same behavioural essentials as the present one, is indicated by the following inter-trial correlations:

MEASURE	r	N
Variability		
(10 and 5 minute trials)	0.710	208
Reaction Speed		
(10 and 1 minute trials)	0.837	316
(10 and 2 minute trials)	0.900	241
(Two 5 minute trials)	0.924	318

Variability was the deviation from the average number of controls pressed for each 30 second period. It was not practicable to obtain a coefficient on the Learning Time section, which would have meant a repetition of the rather lengthy learning procedure on another combination of lights and controls.

Validations of the earlier model yielded the following biserial correlations between test scores and success or failure at Flying School:

MEASURE	r bis	S.E.	N
Learning Time	0.201 0.241	0.089 (Not available)	279 170
Reaction Time	0.363 0.418	0.084 (Not available)	276 170

(Short Learning and Reaction times were favourable to success)

These correlations are somewhat smaller than those obtained on the post-war groups for 1954 or 1955 with the redesigned apparatus. In addition to the fact that the latter was mechanically more efficient than its predecessor, there were some procedural changes in the post-war testing. Greater emphasis was placed on speed, which gave more scope to rapid learning capacity, and by changing the relationships between lights and controls more frequently, there was very little chance that subjects already tested could divulge the combinations to others awaiting their turn.

In 1942 (with the earlier model) an attempt was made to derive an objective temperament measure from the test, on the assumption that Primary-functioning subjects would work more quickly but be more variable from period to period, whereas Secondary-functioning subjects would work more slowly but be more consistent. On a sample of 172, where the subjects had been rated as Primary or Secondary by three or more independent observers, a biserial correlation of 0.48 (± 0.09) was obtained between variability on the test (deviation from the average number of controls pressed for each 30 second period) and the function ratings. The biserial correlation between speed of work and function ratings was 0.30 (± 0.07). A combined speed and variability score gave a biserial correlation of 0.57 (± 0.06). On a later sample of 102, a combined speed and variability score correlated 0.55 (± 0.06) with function ratings. All these results were consonant with the above assumption.

Other research had indicated that, in general, a Primary tendency was desirable for fighter pilots, and Secondary, for bomber pilots. In peace-time, however, with air-training on a very much smaller scale, it was not considered worthwhile to emphasize this distinction at the selection stage.

SECTION I

Summary

1. Constructional details are given of five Co-ordination tests, a Steadiness test, and a complex Hand-foot Reaction test; together with accounts of their application to groups of air-pilot candidates.

2. As indicated by inter-trial correlations, all the tests have useful reliability.

3. When a test such as Handlebars, Variable Co-ordination and Steadiness affords measures of both "CorrectTime" and "Error Time", the former is the more reliable. (In the Roundabout, "Error Time" is quite a different concept as it refers to achievement in a part of the test separate from that in which "CorrectTime" was measured). Where Time and Error are involved in the same activity, as in Handlebars, Variable Co-ordination and Steadiness, there is usually some negative correlation between them, implying that fast working does tend to favour the production of more error. Opposed to this, however, there is some instrumental dependence between the measures whereby fast working reduces error because the subject simply has less time in which to accumulate error. In a complex task like handlebars, which demands very precise and careful work, this effect is less marked, and Time correlates with Error to the comparatively large extent of -0.39. However, in the simpler Steadiness task it is more marked, and while the

average correlation between Time and Error is about -0.10, on some groups there were actually small positive correlations between the measures, indicating that in this test fast working may sometimes be slightly conducive to the production of less error.

4. In all the tests, there are some measures that show a significant relationship to success in flying training. Validity coefficients are biserial correlations between test performance and a Success/Suspension criterion in which suspension was only for reasons of poor practical airmanship:

Floating-effect co-ordination

On two annual groups, overall performance gave respective coefficients of 0.71 and 0.31. The first group, that produced the larger one, was significantly more heterogeneous in its composition.

Handlebars Co-ordination

Only on one of three annual groups tested was there any validity, and this was confined to the first trial : Correct Time, 0.32; Error Time, 0.35. (Slow working and less error in the test, was associated with flying success).

Variable Co-ordination

Results from one group tested indicated that on the "Correct Time" measure three of the four trials were valid : 1, 0.34; 3, 0.42; 4, 0.48. (Fast working in the test was associated with flying success).

Specific Arm-leg Co-ordination

On four of five annual groups tested, total performance gave coefficients ranging from about 0.30 to 0.60. The mean validity coefficient for all five groups (considerably depressed by an inexplicable lack of validity in one group) was 0.31, which is nevertheless highly significant. (In a sixth group, tested subsequently, a significant product moment correlation of 0.40 with practical flying, was obtained). This machine has become a standard "job" test for selecting pupil air-pilots.

Roundabout

Results from one group tested, showed that measures on the Pursuit part of the test ("Error Time") had an aggregate validity of 0.35.

Steadiness

On five consecutive annual groups, the "Correct Time" measure showed small but significant mean validities in two of the four test trials :

1, 0.19; 4 (Stress Trial), 0.22.

(Slow working, probably indicating an absence of "fluster", was associated with flying success).

Hand-foot Reaction

Results from two groups (supported by results from two large war-time groups, obtained with an earlier prototype) indicated that both the initial Learning part of the test and the subsequent Reaction part had useful validity, but the latter, in which the candidate had to apply what he had learnt, quickly and without error, was the more significant, with coefficients ranging from about 0.40 to 0.60.

5. The tests which appear to have the least usefulness as instruments for predicting success in flying training are Handlebars Co-ordination and Steadiness. Possibly the muscular control involved in the former is of a much finer pattern than is ever likely to be required by the air-pilot. This test might give better results in relation to some of the finer industrial occupations. Performance in Steadiness (particularly under stress) is very dependent on personality factors which were conceivably of more importance for success in war-time training, when the test showed useful validity.

Flying, however, remains a basically intimidating vocation, and it is of interest to note that in both Steadiness and Variable Co-ordination the trials that showed the most significant relationship with flying success were those in which there was the threat of punishment by electric shock.

6. It is likely that the validity coefficients obtained with these tests have been depressed by the restriction of range in the criterion groups through pre-selection. This has particular reference to Specific Arm-leg Co-ordination, Steadiness, and Hand-foot Reaction, which formed part of the actual selection battery.

Scores in the other tests did not play any part in the actual selection of the candidates on whom the subsequent validation studies were done, and the depressive effect would therefore have been less in these tests. (See

Section VII, "Reliability and Validity").

7. For various practical reasons, it was not possible to administer all the tests to the same groups of candidates. The few available correlations between tests are mostly rather small and restricted to certain parts of tests, but in the case of Handlebars and Steadiness there appears to be a substantial overall relationship, the coefficients being 0.53 (Time) and 0.40 (Error Time). These and other available correlations are discussed in Section III.

8. Although a group of English-speaking subjects did better than a group of Afrikaans-speaking in certain mental tests, there was no significant difference in Specific Arm-leg Co-ordination. Later results from Floating-effect Co-ordination and Variable Co-ordination (not quoted here) have likewise shown no significant differences between groups.

9. While the tests in this section were intended primarily for air-pilot candidates, some of them might have wider application in both basic research and occupational selection.

SECTION II

SENSORY-MOTOR TESTS INTENDED MAINLY FOR INDUSTRIAL RESEARCH AND CLASSIFICATION

SECTION II

Introduction

Industrial operatives, like air-pilots, are primarily concerned with situations in which practical skills play an important part, but which are, generally less complex, exacting and hazardous than flying situations. In industry, many occupations of the semi-skilled and skilled sort such as tending machines, carrying out mechanical repqirs, handling tools, and assembling components, usually require directed sensory-motor activities that are well controlled and fine or moderately fine, in scope. Others, like the driving of vehicles and winches, may require, in addition to muscular co-ordination, prompt and precise reaction to certain stimuli.

In air-pilot research, there are the problems of identifying and measuring the skills necessary to control one kind of very complex machine. Industrial research is faced with the problems of developing tests that will predict abilities to control many kinds of physical things, which, although generally much simpler than aircraft, often involve patterns of manipulation that are highly specific to themselves.

While it is undoubtedly true that specific "job" tests provide the most adequate means of measuring abilities for practical jobs, it is only in comparatively few specialist vocations that their use is economically justifiable. For the industrial rank-and-file, overall expediency imposes the need to develop instruments having more "general-purpose" applications.

This section comprises three chapters. The first is concerned with two Sensory-motor tests, namely, Chopsticks and Tweezer-Nozzle, in which fine co-ordinated manipulations are performed with implements. The second deals with six Sensory-motor tests that involve moderately fine manipulations, namely, Tweezer-Nuts, Tweezer-Mirror, Two-hand Sticks, Steadiness Stick, Tray Co-ordination and Ring Throwing. Implements are employed in the four former, but not the two latter. In Tweezer-Mirror there is also an element of perceptual-motor conflict.

The third chapter describes five tests of which the first three, namely, Tripod-and-weight, Object Sorting and "Three-dimensional" Peg-board are really hybrids involving both sensory-motor skill and the ability to perceive three-dimensional form relations. For this reason it is preferable to call them "Assembly" tests. Code-signal Reaction, measures choice-reaction to signals given in an elementary code; and Dynamic Coincident Reaction, simple but precise reaction to a moving stimulus.

Reprinted from the Journal of the National Institute for Personnel Research
J. NAT. INST. PERSONNEL RES., 1957, 7, 75–79

TWO TESTS OF IMPLEMENT MANIPULATION.

D. R. De WET.

(Received 5th April, 1957)

Introduction.

Broad distinctions have been mentioned in the past by Muscio (1922) and others between activities which are predominantly mental and those where the elements are mainly motor.

Although it is still a moot point whether the organism can undergo mental experience without some concomitant motor activity, at least on a reflex or autonomic level, there seems to be fairly general agreement that all voluntary motor activity exhibits an essential nexus with some sort of 'mental' process in the central nervous system.

A decapitated snake or fowl is capable of marked motor activity for a short time, but *the movements* are random, without direction or motivation.

Even the simplest manual skills have a 'mental' element which persists in all tests of this type to a greater or less degree despite our attempts to eliminate it. We are still faced with the dilemma which baffled the earlier investigators like Muscio (1922), Perrin (1921), Wyatt (1926) and Seashore (1928), viz. :—

'Mental' or 'Intellectual' tests tend to give significant inter-correlations, whereas 'skill' tests do not.

The 'simpler' the skill tests (i.e. the smaller the 'mental' content), the greater appears their specificity. On the other hand, more complex ones may tend to give higher inter-correlations, but the 'skill' relationship becomes less interpretable on account of increased 'mental' overlap.

The apparent specific nature of 'simple' skill tests is not at present adequate justification for discarding this approach. It may well be that our concept of 'motor skill' is far too general and our interpretation of 'specificity' insufficiently relative.

From the practical standpoint of personnel selection, although the correlations between certain individual skill tests and operational criteria may not be large, their contribution to the battery as a whole can still be of value. In discussing 'Specific duties tests', (which describes the majority of applied skill tests) Biesheuvel (1955) pointed out that : " Even a small gain may be an important consideration in programmes involving scarce or highgrade manpower, costly training procedures or vitally important duties."

Muscio (1922) drew his conclusions from tests as diverse as Aiming, Tapping, Tracing, Form-boards, Putting matches in a matchbox, Wrist movements, Pursuit, Steadiness and Auditory Reaction. In the light of later research it is not surprising that his correlations were low.

The 'Matchbox' test is an example of a 'manipulative dexterity' test as we regard the term today. These seem to fall into a distinct category of motor skill tests, comprising those where the activity is a repetitive process, usually simple, involving a series of separate identical stimuli. Even in this restricted sense, permutations and combinations are many and parallel activities legion. It may be exemplified by an operative placing successive nuts on an assembly job or a simian placing successive nuts (of an edible kind) into his own maw.

Resembling the 'Matchbox' test are certain simple sorting activities and Peg-boards. Most of these have a common feature, viz. the stimuli are directly prehensible to the subject. In another class of manipulation test the stimulus is not handled directly by the subject, instead he manipulates it with a tool or implement of some sort.

The two tests described here are of this latter kind.

Chopsticks Test. (*see figures 1 and 2, pages 77 and* 79.)

This apparatus is intended primarily for individual administration to investigate the more specialized manual skills which appear to prevail in certain occupations, and lends itself also to the study of temperament and personality traits by direct observation of the subject's performance.

Where a measure of straightforward attainment is the main object, several identical models could be administered to a small group at once.

The inclusion of a 'frustration' element is intentional, as this is characteristic of many delicate assembly operations in industrial and other work where a wrong or careless move can suddenly undo the results of much painstaking labour.

Description.

The apparatus consists of a rectangular wooden case having a transparent top and a hole in the middle of each end. The case hinges on a solid wooden base-board designed to rest on a table, and its angle relative to the base-board is made adjustable by means of two slotted struts and thumbscrews. On the bottom of the case is fixed a metal plate perforated with identical holes in rows.

A number of steel balls (corresponding usually to the number of holes in the plate) reposes inside the case.

By manipulating two rods through the holes in the ends of the case, the subject transfers steel balls to the surface of the plate and directs them so that they come to rest over the holes. The score is commonly calculated in terms of the number of balls placed in a given length of time.

The normal angle of the case relative to the board is ten degrees.

To obtain a 'frustration' element, the holes in the plate are made of considerably smaller diameter than the balls, so that although the latter are positively located, they can easily be dislodged by careless movements of the rods.

Technique.

It is customary to tell the subject that he loses credit for the balls which he knocks off; i.e. at the end of the standard time period he scores only according to the balls which are actually on the plate.

As it is somewhat more difficult to get balls placed on the upper rows of holes than on the lower rows, a graded system of scoring may be advisable. If this is decided on the subject had better be informed briefly of the relative values of the rows before he starts the test.

Although the subject is told that he is scored in terms of those balls actually on the plate at the end of the time limit, a great deal of useful information is neglected if we consider only this final picture.

To get the most comprehensive data from this test when it is administered individually, the observer should make a note at short, regular intervals of say a minute or two minutes of the exact number and distribution of balls which the subject has succeeded in locating on the plate.

This can be facilitated by using a scoring sheet having ten or twenty blank charts of the perforated plate printed thereon. Dots may represent positions of holes. At every time interval the observer makes a mark over those dots which correspond to holes where balls are in position.

By this means a continuous and comprehensive record of the subject's progress through the test is obtainable.

Not only will this provide more varied and accurate scope for rating his skill performance, but his individual method and preferences for a particular section of the plate at different stages; the precise locations and occasions of upsetting balls already placed and the possible effects of these mishaps on subsequent performance, are objectively and permanently recorded.

Opposite each chart a space may be provided for the observer to make a quick note of any overt reactions by the subject.

A suitable standard time for normal subjects is fifteen or twenty minutes.

Tweezer-Nozzle Test. (*see figure 3, page 77.*)

Whereas in the Chopsticks test the subject sits comfortably at a table and may steady his arms on its edge, the Tweezer-Nozzle test is intended to investigate manipulative dexterity in relation to a background of wider muscular control and postural dynamics.

As an individual or group test it may be done standing, sitting, or in other positions, provided that the arms are not steadied against the body or any other supports. The object of this precaution is to give hand tremors and general involuntary body movements, unrestricted play.

Description.

There is a metal nozzle, surrounded by a cup-like flange to which it is fixed so that its lower end forming a plug section fits into a hollow handle of bakelite or other suitable material.

The handle is made easily removable from the top section by having two parallel saw cuts in the plug part which give it a measure of lateral spring tension.

It is preferable that the cup and nozzle be made of non-magnetic material.

Technique.

Gramophone needles are placed in the surrounding recess and the activity consists in transferring them by means of a pair of tweezers into the nozzle-tip through which they fall into the hollow handle.

The usual procedure is to distribute 30 needles in more or less random equality in the cup recess and rate performance in terms of the total number of needles transferred into the handle in a standard period of time. For individual application, however, it may be varied. For example, ten needles may be placed in the cup and a measure of time taken in each case to transfer them all, recorded as the score.

Most persons prefer to hold the nozzle in the left hand and the tweezers in the right, but subjects are free to hold them the other way about if desired.

They may insert the needles blunt end or sharp end first at random, but they are instructed not to rest their hands against each other or the tweezers on the cup edge and not to worry about retrieving any needles which may drop on the floor as these are lost points anyway.

Up to a dozen of these tests can be conveniently administered to a group at once.

ACKNOWLEDGEMENTS.

The Chopsticks apparatus was photographed by Mr. J. B. Kirstein.

REFERENCES.

BIESHEUVEL, S. (1955) : The limitations of aptitude tests in selection for specific occupations. *Proc. S. Afr. Psychol. Ass.*

PERRIN, F. A. C. (1921) : An experimental study of motor ability. *J. of Exp. Psychol.* 4, 24-56.

MUSCIO, B. (1922) : Motor capacity with special reference to vocational guidance. *Brit. J. Psychol.*, 13, 57-84.

SEASHORE, R. H. (1928) : The Stanford motor skills unit. *Psychol. Monographs* 39, 51-65.

WYATT, S. (1927) : An experimental study of a repetitive process. *Brit. J. Psychol.*, 17.

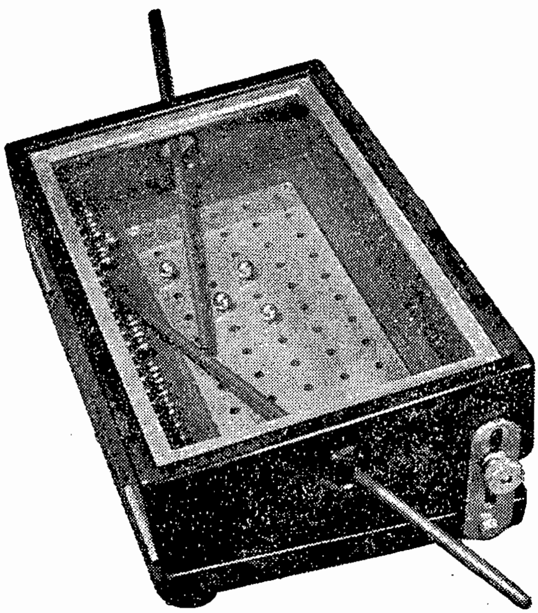


Figure 1.

The Chopsticks Test.

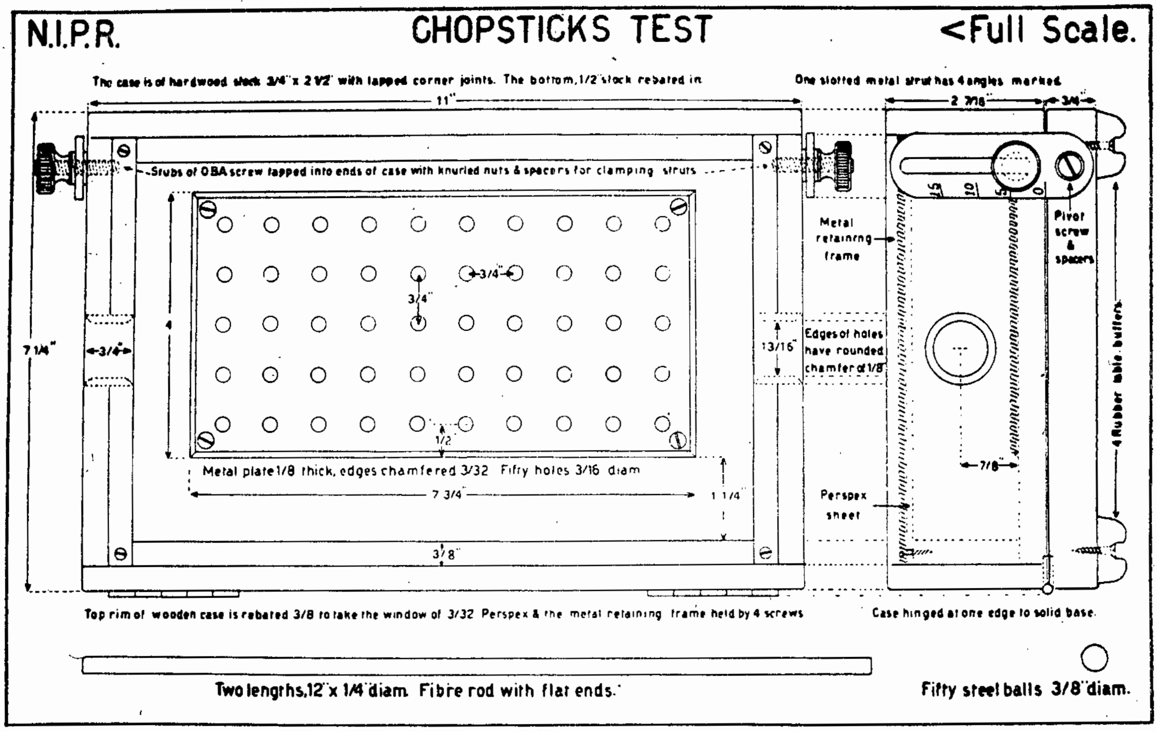


Figure 2.

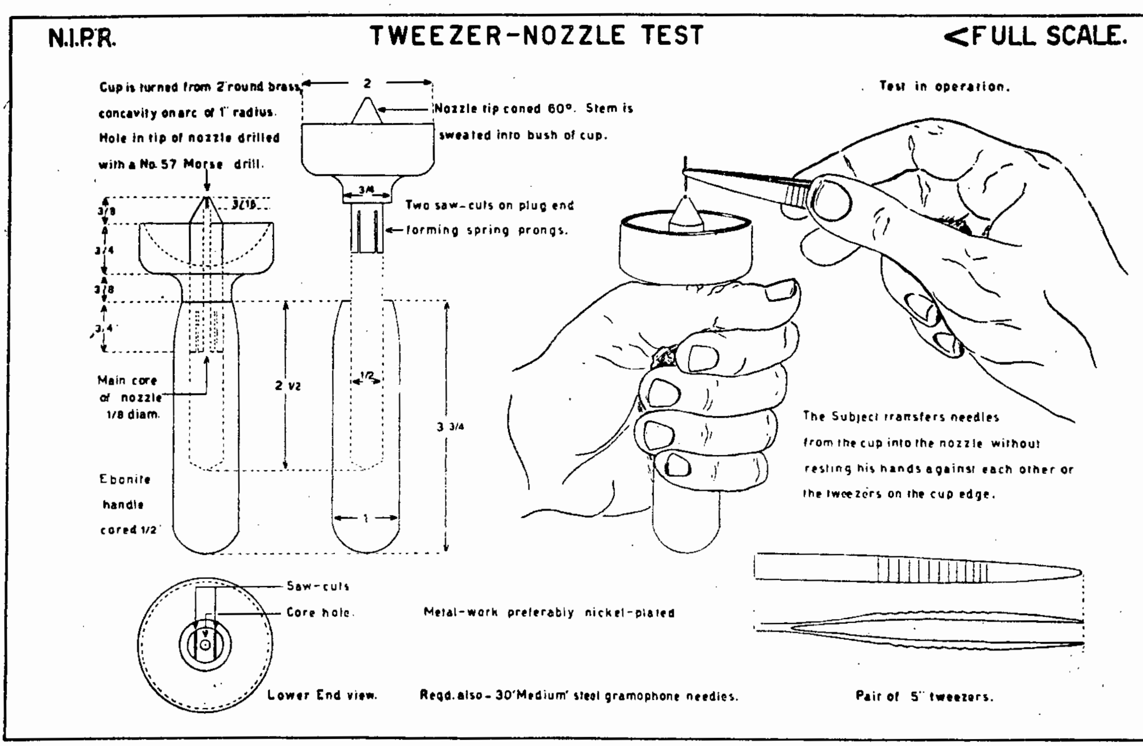


Figure 3.

GOLDFIELDS 18850

APPLICATIONS

African Subjects

The Chopsticks and Tweezer-Nozzle tests were administered by Hudson, Mokoatle and Mbau (2) to 18 untrained African operatives in a factory producing nuts and bolts, with the object of answering the following three questions in relation to the tasks involved in these tests and in operating a "Nutting" machine: (1) Will a period of systematic training even out individual differences in initial ability? (2) Does incentive have any effect on performance? (3) Can the tests predict ability to perform a simple, repetitive operation like "Nutting", which consists essentially of picking up each bolt in turn from a container, holding the threaded end in alighment with a nut rotated by the machine, until the two are joined, and then dropping the assembly into a chute?

Candidates were given five one-minute trials, once only, on Tweezer-Nozzle; two three-minute trials daily, on Chopsticks for 21 days; and two three-minute trials daily on the Nutting machine, for 21 days. On Chopsticks and Nutting, up to the eleventh day, a daily one-minute period of general training in principles of operation preceded the first test trial, and a three-minute period of specific training, the second test trial. Instructional training was then discontinued, as unnecessary, and the stimulus of a small monetary incentive introduced.

Achievement scores on Chopsticks and Tweezer-Nozzle correlated, respectively, 0.59 and 0.51 with achievement scores on Nutting (with incentive), whereas, without the incentive, these correlations were 0.51 and 0.38. Increment scores on Chopsticks and Tweezer Nozzle, reflecting the amount of improvement in performance, correlated 0.59 and 0.04 with achievement scores on Nutting (with incentive). As the test sample is small, only four of these correlations reach the 10% level of significance. In connection with the validities obtained, the investigators state : "Of the two tests which predict Nutting machine achievement score, Tweezer-Nozzle is the more useful. On this test achievement score is the mean of trials 2 - 5. It is possible with a short simple test lasting five minutes and capable of group administration to obtain prediction of eventual work performance after training."

On the twelfth day, the introduction of a small monetary incentive stimulated the Nutting operation, which was a monotonous task having a rather low "status" value in the factory. As performance on Chopsticks had not yet reached a plateau when the incentive was introduced, it was not possible to determine its effect here.

Results obtained from Chopsticks and Nutting over the 21 day period showed that practice and training did not even out individual differences in initial ability. The good starters tended to retain their ascendancy. It was not possible to assess Tweezer-Nozzle in this respect, as it was administered once only.

Chopsticks (Reliability)

Where available, the individual Chopsticks scores of subjects who took part in the above investigation have been examined by the writer with the object of obtaining some indication of range, re-test reliability, and inter-trial consistency.

Table 1 shows the distributions and means of all 18 subjects on the first and seventh days of testing. Scoring was according to the graded system: a ball located in the top row counting 5, and in the consecutive lower rows, 4, 3, 2, 1, respectively. On the first day the distribution was skewed towards the "difficult" side; on the seventh, it was more symmetrical.

Test - Re-test reliability, as indicated by the correlation between performance during seven consecutive days of testing, is shown in Table 2. The coefficient increases from 0.51 on the second and third days, to 0.89 on the fourth, and thereafter remains about this size. Even with such a small sample, the improvement in reliability on the fourth day is significant at better than the 5% level.

Table 3 gives the Mean total score, S.D., and relationship between trials 1 and 2, on each of 19 consecutive testing days. Unfortunately, scores were available on only eleven subjects who were present at every testing session.

The average uncorrected internal consistency of the test was 0.57 for days 1 to 4; 0.85 for days 5 to 7; and 0.85

for days 8 to 11. Maximum consistency was reached by about day 6. The introduction of the monetary incentive to performance on day 12, appears to have had a rather disturbing effect on the internal consistency, which dropped to the lowest level of 0.35, although the overall mean performance was still equal to that of day 11. On day 11, one subject did moderately worse in the second trial than in the first, whereas the rest tended to perform slightly better or about the same in the second trial. On day 12, two subjects showed considerable decrement in the second trial, and two, very great improvement, which in such a small sample depressed the correlation appreciably.

After day 12, when the subjects had become more used to the idea of the monetary incentive, the internal consistency tended to increase again, the mean for days 13 - 15 being 0.69, and for days 16 - 19, 0.71. However, it does not reach the maximum of the pre-incentive period, although the mean performance continued to improve.

While the introduction of the monetary incentive was probably beneficial to all-round achievement, it could have reduced internal consistency because (a) the more strongly motivated subjects sometimes tried modifications in technique, with or without success; and (b) their judgment was sometimes adversely affected by their ambition to excel, and they took risks that resulted in the loss of balls already located.

Chopsticks (European Subjects)

(1)

In an investigation of relationships between stress tolerance and psychological and physiological variates, conducted in the Electro-encephalographic Department of the NIPR, the Chopsticks test was included as a medium for assessing temperamental aspects of behaviour which might be related to differences in the ability to tolerate stress.

The following two sensory-motor tests were also included in the battery :

Visual Selective Reaction

Geometric figures of six different kinds were automatically presented one at a time on a screen, and in each case the subject reacted as quickly as possible by pressing the corresponding one of six keys. If he pressed the wrong key, the figure remained on the screen and an error was recorded. When he pressed the correct key, the figure was replaced by one of the other five at random, or it might be repeated. The scores were : (a) Correct reactions during a continuous period of five minutes. (b) Errors.

Visual Perception Rate

Using a Hodge electronic tachistoscope, eight white cards marked with three to ten black squares were each presented twice in random order for 200 milliseconds. After the exposure of each card the subject had to report the number of dots he perceived, the score being the number of correct responses.

While the subjects performed the Chopsticks test for a continuous period of 10 minutes, temperamental manifestations in terms of Activity, Emotionality and Primary-Secondary Function were observed and rated on scales.

In addition to the psychological tests, which were all performed under normal conditions, a large number of neuro-psychological and physiological measurements were taken, including EEG Alpha Frequency, Heart Rate and Respiration Rate, the two latter under various conditions. All the subjects were European males 18 to 22 years old. Of these, 43 had previously undergone intensive physiological testing at the Human Sciences Laboratory of the Chamber of Mines. Among the many measures taken there on this group, was the subjects' Pyruvic Acid level in millimols after controlled periods of strenuous physical work.

Performance at Chopsticks was according to the graded system, a ball in the top row counting 5, and in the consecutive lower rows 4, 3, 2, 1. The distribution of scores obtained on the sample of 139 is shown in Table 4. It is skewed slightly towards the "difficult" side. The mean score achieved during the continuous 10 minute trial was 34.84 (S.D. 16.11).

Correlations obtained between Chopsticks achievement and some of the other variates, are given in Table 5. There are small positive relationships with Secondary Function (tendency to Cautiousness, Steadiness, Reserve, etc.), Emotional Stability, Visual Selective Reaction (Correct), and Visual Perception Rate; and small negative relationships with Heart Rate (both at rest and during noise), and Respiration Rate (at rest). The noise stimuli were wide spectrum at a level of 92 db, each lasting for 1.75 seconds and presented at short random intervals.

The most significant result is the negative correlation of -0.36 ($P < 0.02$) between Chopsticks Achievement and Pyruvic Acid level after physical work. Mr. G.K. Nelson of the EEG Department considers that biochemical factors influencing behaviour might be responsible for this finding and also for other significant correlations found between Respiration Regularity and Lactic Acid level (-0.31), and GSR Adaptation Rate and excess Lactate (-0.37).

It is not possible to compare the mean Chopsticks scores of the African and European samples, as the former did two three-minute trials, and the latter, a continuous ten-minute trial.

(2)

The Chopsticks test was applied by Holtz (1) in a preliminary analysis to her investigation on temperament and personality factors, with the object of assessing the overt reactions of the subjects in this frustrating manipulative task. Instructions to the subjects (46 European university students) emphasized the skill aspects of the performance, so that they should not suspect the real purpose of the test.

While the subjects worked for fifteen minutes at Chopsticks, they were independently rated by two observers on each of the 24 items of the NIPR Temperament Form, where eighteen items related to the Heymans-Wiersma system, namely, six in each of the categories Primary-Secondary Function, Activity, and Feeling Tone; and the remaining six to Emotional stability and instability. As the ratings of the two observers corresponded closely, they were combined into one. Product-moment intercorrelations were then computed between the 24 items and subjected to factor analysis by Thurstone's method. Five factors emerged, and after rotations directed towards the achievement of simple structure, these were identified by Holtz (11; p.95) as follows:- "(A) Vitality and unrestricted expression of feeling. (B) Emotional disturbance, nervousness and tension. (C) Positive motivation and concentrated effort. (D) Disintegrating and disorganizing aspects of primary function, as evoked by the frustrating Chopsticks test. (E) Agitation; ready mobilization and utilization of energy." Holtz (11; p. 78) remarks : "In our investigation, the Chopsticks test provided a valuable situation for gauging the degree of task involvement and efficiency of motivation."

Tweezer-Nozzle. (Reliability etc.)

Table 6 shows the distributions of total Tweezer-Nozzle scores over five one-minute trials obtained from respective industrial groups of 18 Africans and 18 Europeans, and Table 7, the mean scores achieved by the groups in each trial. In neither range, variability, mean performance in each trial, nor mean overall performance, is there any significant difference between the groups.

The correlations between Trials (2 + 4) and (3 + 5), and the corresponding reliability coefficients obtained by the Spearman-Brown correction, are shown in Table 8. On the present small samples there is no significant difference between the coefficients for the respective groups, although that for the

Europeans is somewhat larger. They indicate that, for either group, a testing period of five one-minute trials has good reliability.

Relationships obtained between the Tweezer-Nozzle test and some other tests administered by Murray (3) to 130 industrial Africans, are given in Table 9, and also the relationship between Tweezer-Nozzle and Chopsticks computed from the data of Hudson et al. (2). It has no significant correlation with Wiggly Blocks and Street Gestalt, which are primarily tests of form recognition, but correlates significantly with Cube Construction, Two-hand Co-ordination (Moede-type) Accuracy, Object Sorting, Screws-and-Nuts Assembly, Letter-and-Numeral Sorting, Tripod Assembly, and Chopsticks. (The six former tests are described in another sub-section pp. 200 - 222). The largest correlations are with Tripod Assembly (0.44; $P < 0.02$) and Chopsticks (0.48; $P < 0.05$), where the patterns of manipulation also require fairly fine muscular control.

Tweezer-Nozzle (Performance and Estimation)

The Tweezer-Nozzle test has been administered under the direction of Dr. Reuning of this Institute, to 193 European mining apprentices ranging in age from 15 to 24 years, with the object of investigating actual performance in relation to subjects' own estimates of their performance. Ten one-minute trials were given, and before starting each trial the subject was asked to "guess" how many needles he would succeed in inserting. The following are some results from a preliminary examination of the test scores by the writer.

Table 10 shows the mean total Performance and Estimation scores, and the mean amount of Error in Estimation in terms of a percentage of the Performance score. The Estimation mean is larger than the Performance mean, the difference being very significant (P = 0.0001) because the data is highly correlated. The mean amount of Error in Estimation by the subjects (irrespective of the direction of the error) was 11.79%.

Intercorrelations between the three measures are given in Table 11. There is no significant relationship between the percentage of Error in Estimation and Estimation. As subjects' estimations were based largely on their actual achievement in preceding test trials, there is a large correlation (0.824) between Performance and Estimation. A significant correlation of -0.350 between Performance and Error in Estimation indicates a tendency for the good performers to predict their achievement more accurately than the inferior performers.

As shown in Table 12, six of the 193 subjects estimated their overall performance exactly; 67 under-estimated by 9.73%; and 120 over-estimated by 13.53%. Error in Over-estimation was significantly larger at better than the 5% level.

Of those subjects who estimated either over or under, significantly more tended to over-estimate, the 99% confidence interval for the proportion of those who under-estimated being 0.275 to 0.450, which shows that the value 67/187(0.36) is significantly different from 0.50 at something beyond the 1% level.

Tweezer-Nozzle (Effect of Alcohol)

In an investigation on the behavioural effects of mild to moderate alcoholic intoxication, Nelson (4) applied various EEG measures and the following sensory-motor tests, to 29 African male adults :

(1) Tweezer-Nozzle; (2) Improved Steadiness Apparatus (p); (3) Code-Signal Reaction (p); (4) Simple Reaction Time; (5) Two-hand Co-ordination (Moede-type); (6) Visual Perception Rate; (7) Visual Discrimination; (8) Tapping Speed; (9) Two-hand Pursuit; (10) Card-Sorting.

The only sensory-motor tests in which performance was significantly inferior under conditions of mild to moderate intoxication were : Visual Discrimination, Tapping Speed, Card Sorting, and Two-hand Co-ordination (Accuracy).

ACKNOWLEDGMENTS

I am grateful to Miss Ina van Wyk, Mr. G.K. Nelson, Mr. C.O. Murray, Mr. B. Mokoatle, Mr. G.G. Mbau, Dr. W. Hudson and Dr. H. Reuning, for experimental data.

REFERENCES

1. HOLTZ, Shirley. A Factorial Analysis of Temperament and Personality Aspects with Special Reference to Emotional Adjustment. M.A. thesis. Johannesburg, Univ. of the Witwatersrand, 1959.

2. HUDSON, W., B. MOKOATLE and G.G.MBAU. The influence of training and practice in the test and work performance of a small sample of African workers. J. Nat. Inst. Personnel Res., 1958, 7, 88 - 94.

3. MURRAY, C.O. The Structure of African Intelligence. A Factorial Study of the Abilities of Africans. M.A. thesis. Durban, Univ. of Natal, 1956.

4. NELSON, G.K. A Neuropsychological Study of the Comparative Intoxicating Effects of Various Alcoholic Beverages. National Institute for Personnel Research, Contract Report G30/62, Johannesburg. Nov. 1962.

TABLE I

Chopsticks

(African subjects. Two 3 minute trials)

1st Day		7th Day	
Scores	N	Scores	N
0 - 4	3	29 - 35	1
5 - 9	2	36 - 42	0
10 - 14	1	43 - 49	2
15 - 19	2	50 - 56	3
20 - 24	4	57 - 63	1
25 - 29	2	64 - 70	1
30 - 34	1	71 - 77	1
35 - 39	1	78 - 84	4
40 - 44	0	85 - 91	2
45 - 49	0	92 - 98	1
50 - 54	1	99 - 105	0
55 - 59	1	106 - 112	2
	18		18
Range : 0 - 56 Mean : 21.61 S.D. : 15.36		Range : 29 - 109 Mean : 71.83 S.D. : 21.42	

TABLE 2

Chopsticks

Test - re-test Reliability

(African subjects N = 16)

Days	r
1 and 2	0.51
2 and 3	0.51
3 and 4	0.89
4 and 5	0.86
5 and 6	0.89
6 and 7	0.89

TABLE 3

<u>Chopsticks</u>

Means and S.D's (Trials 1 & 2), and correlation between trials 1 and 2, on successive days of testing (African subjects N = 11)

Day	Mean	S.D.	r	Rel.
1	20	16	.48	.65
2	35	17	.63	.77
3	41	17	.59	.74
4	50	20	.56	.72
5	53	25	.79	.88
6	62	29	.90	.95
7	67	24	.85	.92
8	73	27	.94	.97
9	75	24	.69	.82
10	87	29	.89	.94
11	92	32	.88	.94
12	91	24	.35	.52
13	92	24	.78	.88
14	101	31	.74	.85
15	96	24	.56	.72
16	105	32	.79	.88
17	89	21	.46	.63
18	110	36	.82	.90
19	104	35	.76	.86

TABLE 4

Chopsticks

(European subjects. One 10 minute trial)

Scores		N
5 - 10		7
11 - 16		11
17 - 22		16
23 - 28		22
29 - 34		16
35 - 40		18
41 - 46		18
47 - 52		10
53 - 58		10
59 - 64		7
65 - 70		3
Above	(Scored 101)	1
		139

Range : 5 - 101

Mean : 34.84 S.D. : 16.11

TABLE 5

Chopsticks

Correlations with other variates

(Decimal points omitted)

Other variates	r	N
Primary Function	02	133
Secondary Function	18s	121
Activity	17	121
Emotional Stability	23s	117
Visual Selective Reaction (Correct)	22s	139
Visual Selective Reaction (Errors)	-11	139
Visual Perception Rate	19s	139
Hours slept previous night	06	119
Heart Rate (Rest)	-19s	139
Heart Rate (During noise)	-20s	139
Heart Rate (after noise)	-13	139
Respiration Rate (Rest)	-19s	139
Respiration Rate (During noise)	-16	139
Respiration Rate (after noise)	-07	139
E.E.G. Alpha Frequency	08	139
Age	15	139
Pyruvic Acid level after work	-36s	43

"s" implies significance at the 5% level or better.

TABLE 6

Tweezer-Nozzle

Total scores of five one-minute trials

Scores	Africans		Europeans	
29 - 33	1	Range :	1	Range:
34 - 38	4	29 - 83	2	31 - 86
39 - 43	1		1	
44 - 48	0		2	
49 - 53	4	Mean :	1	Mean :
54 - 58	1	54.22	3	57.00
59 - 63	2		4	
64 - 68	0		0	
69 - 73	2	S.D. :	1	S.D. :
74 - 78	2	15.75	0	16.03
79 - 83	1		0	
84 - 88	0		3	
Total N	18		18	

TABLE 7

Tweezer-Nozzle

Means and S.D.'s of five one-minute trials
and total

Trial No.	Africans N = 18		Europeans N = 18	
	Mean	S.D.	Mean	S.D.
1	8.56	3.06	10.17	3.30
2	10.94	4.85	11.17	3.70
3	11.39	4.26	12.28	4.19
4	11.67	3.21	11.50	2.95
5	11.67	3.28	11.89	4.39
Total	54.22	15.75	57.00	16.03

No significant differences between groups

TABLE 8

Tweezer-Nozzle

Correlations between Trials (2+4) and (3+5)

Group	N	r	Rel.
African	18	0.79	0.88
European	18	0.91	0.95

TABLE 9

Tweezer-Nozzle

Correlations with other tests (Africans)

(Decimal points omitted)

N = 130	r	P
Wiggly Blocks	10	> 23
Street Gestalt	16	> 07
Cube Construction	28	< 02
Two-hand Co-Ord.(Acc.)	28	< 02
Object Sorting	29	< 02
Screws-and-nuts Assembly	31	< 02
Letter-and-Numeral Sort.	32	< 02
Tripod Assembly	44	< 02
N = 18	**r**	**P**
Chopsticks	48	< 05

TABLE 10

<u>Tweezer-Nozzle</u>

Means and S.D.'s (Europeans)

(Totals of 10 one-minute trials)

N = 193	Mean	S.D.
Performance	130.06	30.69
Estimation	135.34	32,20
% Error in Estimation	11.79	12.91

TABLE 11

Tweezer-Nozzle

Intercorrelations (N = 193 Europeans)

	r	P
Performance and Estimation	0.824	< 0.0001
Performance and Error in Estimation (Difference / Performance)	-0.350	< 0.001
Estimation and Error in Estimation (Difference / Performance)	-0.109	> 0.110

TABLE 12

Tweezer-Nozzle

Direction and Accuracy of Estimation

(N = 193 Europeans)

Estimation	N	Percent Error	S.D.	P of diff. Error M's
Over	120	13.53	14.16	< 0.05
Under	67	9.73	9.96	
Correct	6	0.00	-	-

Most subjects ($P < 0.01$) tended to over-estimate.

SIMPLE SKILL TESTS APPLIED TO AFRICANS

ABSTRACT

The construction and experimental application of some "simple" tests of manipulative skill, suitable for administration to several subjects at a time, is described. While using the minimum of apparatus, the battery was intended to cover factors such as dexterity, steadiness, aiming, co-ordination and perceptual-motor conflict. Results obtained on adult Africans indicated that: (1) Reliability of most of the tests was reasonably good. (2) Educational level correlated negatively with only one sensory-motor test, namely, Steadiness. (3) A practical mental test correlated with tweezer-manipulation by mirror-view. (4) The mean intercorrelating between all the sensory-motor tests was about 0.15. (5) In a less intelligent male group the mental test showed a mean overlap of 0.24 with the sensory-motor tests; in a more intelligent female group the overlap was 0.12. (6) There were some differences in achievement attributable to differences in sex and age. (7) The tests showed no useful potentiality for selecting operatives in a porcelain factory.

Two principal drawbacks of apparatus tests, as compared with "pencil-and-paper" tests, are the costliness of reproduction in quantity and the general unsuitability for administration to more than one subject at a time. However, in studying practical skills the use of "three-dimensional" apparatus is often necessary and for certain basic research, or the selection of specialist, high-grade personnel, individual testing on expensive equipment is usually justified; but this is not worthwhile in many elementary investigations, or for low-level industrial selection. Here, the equipment should be constructionally simple, compact to store and transport, cheap to reproduce, and suitable for administration to several subjects at a time.

The following simple pieces of apparatus, used in certain combinations, provide a battery of six or more tests of manipulative skill intended to cover several factors such as dexterity, steadiness, aiming, co-ordination and perceptual-motor conflict. The tests are well suited to illiterates as most of the instruction is by practical demonstration. Up to six subjects at a time can be conveniently managed by one observer.

Description of Apparatus

A basic component in five of the tests is a small square pan with a stainless steel rod mounted vertically at the middle, as shown in Plate 1. For convenience in packing, the rod is detachable from a fixed bush in which it fits snugly, and to afford some means of adjustment to compensate for wear which would tend to loosen the fit, the lower end of the rod is split. When the rod is in position there should be no perceptible wobble at the free end. This end is flat, but the sharp circumferential edge is rounded very slightly with a fine file. The pan is made of a fairly heavy gauge of iron sheet because the extra weight improves stability in use and there is also no need to solder the corners. The tweezers are of a standard round-ended surgical pattern set with a gap of approximately three-quarters of an inch between the ends at the open position.

The other implements are shown in Plate 2. Both metal pins in the Steadiness-stick are a light driving fit through the wood. The ends of the cross-pin may be round or chamfered, but the protecting end of the other pin, like that of the rod in Plate 1, is flat, with the edge rounded only very slightly. Similarly, the ends of the wooden Two-hand sticks are flat with edges rounded just a little. The wooden ferrule at the middle of each of these sticks is glued fast. The rubber rings (Plate 2) are of a standard fruit-jar type readily obtainable in packets of a dozen.

Plate 3 illustrates a simple case bent from a single piece of iron sheet to provide support for a mirror at one end and a screen at the other. The mirror, permanently fixed in a light tin-plate sheath to protect its backing, is held in two channels formed by bending over the vertical edges of the support, and it can be slid out easily for transportation or storage.

The tin-plate tray depicted in Plate 4 has a ramp along the middle of the large compartment that slopes up to the partition and the narrow opening in a tunnel communicating with the small compartment. An upright baffle is soldered to the surface of the ramp directly opposite the narrow opening. The "Perspex" cover for the tray is a friction fit in the rebates so that it does not slide out during use and is nevertheless easily removable for occasional dusting of the interior or to take the balls out to prevent rattling when the apparatus is transported long distances.

Both the Mirror case (Plate 3) and the Co-ordination tray (Plate 4) are spray-painted a dark grey, but it is important that there be no small lumps or other irregularities in the paint surface on the ramp in the latter.

Procedures

Tweezer Nuts (Plate 5, Upper)

The small square pan, with the vertical rod in position and twenty nuts lying flat and randomly about its base, is set on the table within easy reach of the sitting subject. The observer demonstrates, how, when the signal to start is given, the nuts must be picked up one at a time with a pair of tweezers and slipped over the rod. There are two trials of half-a-minute each. After a trial, the number of nuts placed on the rod is recorded and the nuts are dropped back randomly in the pan, ensuring, however, that they all lie flat.

Tweezer Mirror (Plate 5, Middle)

The case holding the mirror is set on the table within easy reach of the sitting subject, so that the mirror faces him, and the pan with the vertical rod and nuts is put inside the case against the upright part of the screen and laterally more or less at its middle. The subject performs the same task as in the preceding test, but is now permitted only a mirrored view of his activity. There are two trials of four minutes each.

Two-hand Sticks (Plate 5, Lower)

The pan with the vertical rod and nuts is placed on the table within easy reach of the sitting subject, and he is required to transfer one nut at a time to the rod, using the two short sticks, one in either hand. No part of the subject's hand may extend beyond the shoulder of the ferrule at the middle of the stick. There are two trials of two minutes each.

Steadiness Stick (Plate 6, Upper)

The subject sits upright and stretches out his arm and hand over the table. The observer then sets the pan and rod on the table in line with the subject's arm at a distance of half the measuring stick (7½") from his finger tips to the rod. Using the Steadiness stick, held in one hand behind the cross-pin, the subject is required to pick up nuts one at a time with the pin at the other end and transfer them to the rod. He may not rest his arms or the stick on the table. There are two trials of three minutes each.

Tray Co-ordination (Plate 6, Middle)

The tray, having the ten steel balls in the large compartment, is held by the subject in both hands so that the smaller compartment is away from him. By suitable tilting the balls must be guided carefully up the ramp, past the baffle, and through the opening into the small compartment. It is not permissible to try flipping the balls directly on to the higher end of the ramp near the opening. There are two trials of four minutes each. After a trial the number of balls in the small compartment is recorded, and this end of the tray is then raised and turned over by the observer to roll the balls back into the large compartment.

Ring Throwing (Plate 6, Lower)

The subject sits upright and stretches out his arm and hand over the table. The observer then puts the pan and rod on the table in line with the subject's arm and at a distance of a full measuring stick (15") from his finger-tips to the rod. The mirror case is set up on end about six inches behind the edge of the pan so that its bottom faces the subject and affords a uniform background and a convenient stop for the thrown rings. (Mr. Currie made this useful suggestion)

A bundle of twelve rubber rings is given to the subject, who is required, without leaning forward, to throw one ring at a time over the rod. Twelve rings constitute a trial, and there are ten trials in all. After each, the number of rings correctly thrown over the rod is recorded. There is no time limit.

In all the above tests, the observer, relying mainly on mime, demonstrates exactly what the subjects have to do, but does not allow them any preliminary practising. A gong-stroke or other suitable sound signals the beginning and end of each trial. In the one-hand tests, the preferred hand is used. In all tests, except the Ring Throwing, subjects are instructed to work as quickly as they can.

Form Series Test (Not illustrated)

This is primarily a test of inductive reasoning ability administered through a practical medium, the instruction being by both verbal explanation and mime. The subject is provided with a board down the left-hand side of which are shown incomplete sections of a recurrent pattern of geometrical figures that may vary in shape (triangle, square or circle), size (large, medium or small) and colour (red, blue or yellow). His task is to complete the pattern in each row by adding two appropriate plastic figures selected from a quantity randomly mixed in a tray. The first five rows are practice items, and the subsequent eighteen, test items, in which the patterns become progressively more complicated. There is no time limit. A point is scored for every item-row completed correctly. Full details of the test are given by Grant (2) who developed it.

Application and Results

The tests were administered experimentally to African employees of both sexes at a newly-started factory for producing ceramic-ware, the objects being to investigate: (a) Reliability; (b) Validity for occupational selection; (c) Intercorrelations between the tests; (d) Relationship to educational level and intelligence; and (e) Sex, tribal and age differences in test performance.

The inter-trial correlations of the six sensory-motor tests and the corresponding reliability coefficients obtained by the Spearman-Brown formula for "split-half" correlation, are given in Table 1. All the tests appear to have acceptable reliability, the worst being Tweezer Nuts, and the best, Ring Throwing. Tweezer Nuts was both the shortest test and the first to be administered; the novelty of the test situation could have induced temperamental and other extraneous influences. It is noteworthy that the coefficient for the female group is considerably the smaller, although on the present restricted samples the difference is not statistically significant. Ring Throwing, in addition to being performed as the last test in the battery, after subjects had got well used to the testing situation, was also the longest with many more trials than the others. The reliability of Tweezer Nuts could probably be improved by increasing the number of trials to four.

Table 2 shows the respective mean scores and standard deviations of the male and female groups, and also the critical ratios of the differences between the means. On educational Level and Form Series (the mental test) the female group was markedly superior. On Tweezer Nuts the male group was slightly superior. The pattern of movements in this task bears a certain resemblance to those in using small hand-tools like pliers and spanners and the males might have had the advantage of more previous experience at this kind of manipulation. It is also possible that the females were more emotionally affected by the strangeness of the initial test situation. On Tweezer Mirror, the male group also performed slightly better, but the significance of the difference does not reach the 5% level. Although this test introduced a perceptual set alien to that in Tweezer Nuts because of the fore-and-aft reversal of direction in the reflected image, the motor patterns in both were much the same.

On Two-hand Sticks, Steadiness Stick, Tray Co-ordination and Ring throwing, there are no differences in the mean scores. As a broad tendency, the male group shows more heterogeneity (reflected by larger S.D's) on all the variates except Steadiness Stick, but considered individually most of these differences are not significant.

Tables 3 and 4 show the intercorrelations between tests in the male and female groups. In both, Educational Level correlates with the mental test Form Series, but is negatively related to Steadiness Stick. The more educated Africans were those who had the mixed blessing of more contact with European culture, many aspects of which, markedly in urban areas, tend to promote neural and cardio-vascular conditions unfavourable to "steadiness". There is also the likelihood that the less-educated "tribal" Africans had more experience with the manipulation of a longish stick both as a weapon and a tool.

Form Series is positively related to Tweezer Mirror, which, like conventional mirror-tracing tasks, particularly when applied to young or other wise low-grade subjects, probably has a moderate loading of g. Because of the extra dimension involved in Tweezer Mirror it is a rather more frustrating task than mirror-tracing on a flat surface. In the intellectually inferior male group, mental factors would have played a greater part in the mirror task, and the larger correlation with Form Series is in the expected direction, although, on the present small samples, the difference in size between the respective male and female co-efficients is not statistically significant.

Steadiness Stick correlates with Two-hand Sticks. Although one requires two hands and the other, one only, both tasks involve the remote control of objects by "stick" manipulation.

In the intellectually inferior male group, Ring Throwing and Tray Co-ordination correlate with Educational Level, and Tray Co-ordination also with Form Series. In the female group, Form Series is related to Tweezer Nuts, in which the female mean was lower. The more intelligent females did better at a sensory-motor task which was generally rather more difficult for females, whether through lack of previous tool-wielding experience or their greater sensitivity to the novelty of the test situation.

In the male group, the most specific tests were Tweezer Nuts, (mean r, 0.03) and Ring throwing (mean r, 0.09); and in the female group, Ring Throwing (mean r, 0.05) and Tray Co-ordination (mean r, 0.07). While Tweezer Nuts, at which the males performed better, shows a high degree of specificity in the male group, in the female group it is the test which shows the widest overlap with others (mean r, 0.25). Among the males, the test showing the widest overlap with others is Tweezer Mirror (mean r, 0.30). In both groups, Tray Co-ordination is related to only one other sensory-motor test: Tweezer Mirror (males), and Tweezer Nuts (females).

In the intellectually inferior male group, Form Series (the mental test) has a mean correlation of 0.24 with the sensory-motor tests; in the superior female group, the corresponding mean correlation is 0.12. This fits in with general findings that the lower the intellectual grade of the subjects, the more significant becomes the rôle of intelligence in performing various tasks.

In both the male and female groups, the mean intercorrela-tion between all the sensory-motor tests is of the order of 0.15.

Validity

To obtain some indication of the usefulness of the tests for predicting success at the factory jobs, the mean test scores of 25 operatives who rated high accoding to observation of their manual proficiency on the job, were compared with the means of 25 who rated low. The results are given in Table 5.

Higher educational level is associated with higher job-rating at between the 10% and 5% level of significance. On Steadiness Stick, the group rated low at their jobs performed better than the one rated high, the difference being significant at about the 5% level. This appears to be consistent with the finding that there was a negative correlation between Educational Level and Steadiness Stick. None of the other tests distinguishes either way between the job groups.

It should be mentioned that much of the testing was done under very makeshift conditions before the factory was in production, and most of the operatives had been at their jobs for only a short time when the subsequent merit ratings were carried out. However, none of the tests was specifically designed to simulate the functional patterns of any of the factory jobs, and, as Vernon (3, p.103) has observed, ..."the notion of dexterity or handiness as a general factor in manual occupations, which can be measured by one or two pegboard, nut and bolt or similar tests, should be discouraged".

Tribal Origin

The means and standard deviations of four tribal groups and the significance of the differences between means and between S.D.'s are given in Tables 6 and 7. Tribal classification of the combined male and female samples was according to general linguistic and cultural affinity: (1) Zulu, Xhosa and Swazi; (2) Sotho and Tswana; (3) Pedi and Ndebele; (4) Shangaan, Tsonga and Venda.

None of the Chi-squared values in Table 7 derived by Bartlett's analysis for homogeneity of variance (1) is as large as 7.81, which is necessary for the 5% level of significance with three degrees of freedom. Hence the differences between the S.D.'s are not significant. As none of the F-values in the right-hand column obtained by an analysis of variance for homogeneity of means reaches the minimum 2.68 required here, there are no significant differences between the means. It would appear, therefore, that in the present tests the tribal origin of the subjects had no influence on their achievement.

Age Differences

The means and standard deviations of four age groups, and the significance of the differences between means and between S.D.'s are shown in Tables 8 and 9. All subjects, irrespective of sex or tribal origin, were classified according to age into the following groups: 1(16 to 20), 2(20 to 25), 3(26 to 35), and 4(36 to 55). By Bartlett's Chi-squared analysis for homogeneity of variance, none of the differences between the S.D.'s (Table 8) reaches the 5% level of significance, as none of the

Chi-squared values (Table 9) is as large as 7.81. The analysis of variance for homogeneity of means produced the F-values in the right-hand column of Table 9. Four of these, which exceed 2.68, indicate that there are differences at better than the 5% level on Educational Level, Form Series Test, Tweezer Nuts and Tray Co-ordination.

With regard to Educational Level (Table 8), the two junior groups 1 and 2 (which are not significantly different) are the highest, Group 3 is next, and Group 4 the lowest. The critical ratios between means adjacent in size are: 1 and 2, 1.50; 2 and 3, 3.30; 3 and 4, 2.05. On Form Series, the two junior Groups 1 and 2 are superior to the two senior Groups 3 and 4, but the test does not discriminate between Groups 1 and 2 or 3 and 4, the critical ratios between adjacent means being 1 and 2, 0.75; 2 and 3, 2.53; 3 and 4, 0.52.

On Tweezer Nuts, Group 1 is inferior to all the others, the critical ratios being, 1 and 2, 2.00; 2 and 4, 0.79; 4 and 3, 0.10. Possible explanations for this are: (a) The most immature subjects were most adversely affected by the novelty of the test situation. (b) The majority of the very youngest were females, who, having had little previous experience in manipulating hand-tools would have found this initial test particularly strange. On Tray Co-ordination, the critical ratios are as follows: 1 and 4, 0.77; 4 and 3, 0.83; 3 and 2, 1.40; (1 and 3, 1.67) none of which is significant. However, there is a significant difference between Group 1 and Group 2, and probably also between Group 4 and Group 2. It would seem that Group 1 (the most junior) and Group 4 (the most senior) did not fare as well as Group 2.

Summary and Conclusions

(1) All the sensory-motor tests appear to be acceptably reliable. Tweezer Nuts, the initial one, and the shortest, showed the lowest reliability, and Ring Throwing, which was both the longest and the last test, the highest.

(2) Females were superior on Educational Level and Form Series (Mental test), and males on Tweezer Nuts (also to a slight extent on Tweezer Mirror), where previous experience with small hand-tools might have been an advantage.

(3) In both male and female groups, Educational Level correlated positively with Form Series and negatively with Steadiness Stick. The more educated Africans were possibly less familiar with a stick as a tool, or just less steady because of closer association with certain enervating influences in European culture. Form Series related to Tweezer Mirror, probably because of an overlap of g. Steadiness Stick and Two-hand Sticks correlated. Ring Throwing was a rather specific test. For either males or females, the mean intercorrelation between all the sensory-motor tests was of the order of 0.15.

(4) In the less intelligent male group, Tweezer Mirror showed the largest overlap with other sensory-motor tests (mean r, 0.30). Form Series (mental test) correlated with Tray Co-ordination and gave a mean correlation of 0.24 with the sensory-motor tests. In the female group, Tweezer Nuts

showed the largest overlap with other sensory-motor tests (mean r, 0.25) and correlated 0.30 with Form Series (mental test). In this more intelligent group, Form Series gave a mean correlation of 0.12 with the sensory-motor tests.

(5) Younger subjects tended to have a higher educational level and also to perform better at Form Series. The very young ones, who were mainly females, fared the worst at Tweezer Nuts, (the initial sensory-motor test). On Tray Co-ordination the most junior group, and to a lesser extent, the most senior, scored less than an intermediate group.

(6) The tests, which were not specifically designed to simulate or any of the manual occupations in a factory producing ceramic-ware, showed no validity with a criterion of merit ratings in these occupations.

(7) The simple design of mirror and case described here can be used in conjunction with various other small manual activities, including conventional mirror-drawing on paper sheets. Two-hand Sticks is particularly interesting when performed by mirror view, but because of its difficulty this combination was not applied in the present study.

ACKNOWLEDGEMENTS

I am grateful to: Mr. H.H.K. Currie, Mr. O. Krenzer and Mr. S. Mngomezulu for test administration; Miss C.M. Elder for computational assistance; and Mr. R.S. Hall for statistical consultation.

REFERENCES

1. BROWNLEE, K.A. Industrial Experimentation, 4th ed., pp.38-40. London, H.M. Stationery Office, 1949.

2. GRANT, G.V. The Construction of a Non-verbal Test of Reasoning Ability for African Industrial Workers. M.A. thesis. Johannesburg, Univ. of the Witwatersrand, 1965.

3. VERNON, P.E. The Structure of Human Abilities. London, Methuen, 1950. 2nd ed. 1961.

TABLE 1

Inter-trial correlations and reliability coefficients

of Sensory-motor tests

(Decimal points omitted)

Sensory-motor Tests	Males 55		Females 58		Comb. 113	
	r	Rel.	r	Rel.	r	Rel.
Tweezer Nuts	60	75	42	59	59	74
Tweezer Mirror	65	79	64	78	65	79
Two-hand Sticks	64	78	63	77	66	80
Steadiness Stick	82	90	75	86	81	90
Tray Co-ordination	63	77	70	82	70	82
Ring Throwing	86	92	84	91	90	95

TABLE 2

Means and standard deviations of male and female subjects

	Males N = 63		Females N = 62		C.R. Diff. Means
	Mean	S.D.	Mean	S.D.	
Educational Level	2.56	2.82	4.77	1.79	5.27
Form Series Test	7.70	5.28	12.03	3.83	5.26
Tweezer Nuts	22.48	6.08	19.98	5.76	2.35
Tweezer Mirror	8.46	6.84	6.18	6.38	1.93
Two-hand Sticks	13.46	7.34	12.13	7.10	1.03
Steadiness Stick	15.14	8.76	16.94	9.53	1.09
Tray Co-ordination	11.94	6.11	11.08	6.03	0.79
Ring Throwing	80.84	18.84	84.66	18.04	1.16

TABLE 3

Intercorrelations between tests ofi Male group N = 63

(Decimal points omitted)

Tests	1	2	3	4	5	6	7	8
1 Educ. Level								
2 Form Series	33s							
3 Tweezer Nuts	07	09						
4 Tweezer Mirror	06	46s	04					
5 Two-hand Sticks	11	22	-07	48s				
6 Stead. Stick	-24s	12	21	32s	34s			
7 Tray Co-ord.	33s	30s	22	37s	14	-05		
8 Ring Throw.	27s	22	-26	27s	27s	14	02	
Mean r's of Sensory-motor tests			03	30s	23s	19s	14s	09

"s" implies significance at the 5% level or better

TABLE 4

Intercorrelations between tests on Female group N = 62

(Decimal points omitted)

Tests	1	2	3	4	5	6	7	8
1 Educ. Level								
2 Form Series	52s							
3 Tweezer Nuts	12	30s						
4 Tweezer Mirror	09	29s	32s					
5 Two-Hand Sticks	08	01	33s	13				
6 Stead. Stick	-26s	-02	27s	06	38s			
7 Tray Co-ord.	-06	11	31s	12	-06	-02		
8 Ring Trow	-05	03	01	12	-02	12	00	
Mean r's of Sensory-motor tests			25s	15s	15s	16s	07	05

"s" implies significance at the 5% level or better

TABLE 5

Means and standard deviations of low-rated and high-rated groups of factory operatives.

Tests	Low N = 25		High N = 25		C.R. of Diff. Means
	Mean	S.D.	Mean	S.D.	
Educational Level	3.00	2.55	4.32	2.36	1.90
Form Series Test	10.64	5.03	11.56	4.64	0.67
Tweezer Nuts	22.84	4.65	23.00	7.53	0.09
Tweezer Mirror	7.64	6.92	8.32	6.82	0.35
Two-hand Sticks	13.40	8.15	10.96	6.33	1.18
Steadiness Stick	17.60	11.44	12.16	7.39	2.00
Tray Co-ordination	13.48	5.49	11.80	6.14	1.02
Ring Throwing	83.20	17.97	78.96	17.68	0.84

TABLE 6

Means and Standard deviations of four tribal groups

Tests	1. N = 15		2. N = 41		3. N = 40		4. N = 28	
	Zulu, Xhosa, Swazi.		Sotho, Tswana.		Pedi, NdeBele.		Shangaan, Tsonga, Venda.	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Education Level	3.69	3.09	3.88	2.47	3.95	2.62	2.96	2.38
Form Series Test	8.75	5.85	10.90	4.74	10.43	5.00	8.64	5.07
Tweezer Nuts	19.80	4.55	21.51	5.51	20.98	6.41	21.75	6.98
Tweezer Mirror	5.33	5.98	7.95	7.71	8.40	6.69	6.14	5.18
Two-hand Sticks	9.47	5.01	13.20	7.81	12.93	7.34	13.57	7.02
Steadiness Stick	11.53	8.12	15.81	9.08	16.60	9.41	18.04	9.20
Tray Co-ordination	10.60	6.63	11.76	6.09	12.58	6.00	9.96	5.78
Ring Throwing	77.80	24.07	84.32	17.27	83.18	17.68	83.32	18.23

TABLE 7

Significance of the differences between the means and S.D.'s of the tribal groups

Tests	Chi-Squared	F Value
Educational Level	1.59	1.05
Form Series Test	1.02	1.54
Tweezer Nuts	3.97	2.52
Tweezer Mirror	5.04	1.20
Two-hand Sticks	3.50	1.22
Steadiness Stick	0.42	1.73
Tray Co-ordination	0.36	1.15
Ring Throwing	2.81	0.47

TABLE 8

Means and Standard deviations of four age groups

Test	1. N = 21		2. N = 57		3. N = 37		4. N = 10	
	16 to 20 yrs.		20 to 25 yrs.		26 to 35 yrs.		36 to 55 yrs.	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Educational Level	5.10	1.84	4.35	2.27	2.58	2.74	1.10	1.79
Form Series Test	11.71	3.84	10.93	4.73	8.16	5.58	7.20	5.07
Tweezer Nuts	18.24	5.52	21.19	6.43	22.68	5.58	22.50	4.48
Tweezer Mirror	6.67	6.98	7.51	6.59	7.32	7.17	7.70	5.52
Two-hand Sticks	12.86	7.00	12.67	7.42	13.43	7.75	11.10	4.68
Steadiness Stick	11.76	7.92	17.23	9.47	16.24	8.47	17.40	10.83
Tray Co-ordination	8.48	6.01	13.07	5.82	11.27	6.29	9.90	4.10
Ring Throwing	83.48	15.51	85.49	19.07	79.00	18.77	79.30	19.35

TABLE 9

Significance of the differences between the means and S.D.'s of the age groups

Tests	Chi-Squared	F Value
Educational Level	5.07	11.08_s
Form Series Test	3.49	4.43_s
Tweezer Nuts	2.36	2.69_s
Tweezer Mirror	1.01	0.09
Two-Hand Sticks	3.01	0.28
Steadiness Stick	1.81	1.97
Tray Co-ordination	2.21	3.46_s
Ring Throwing	1.24	1.06

"s" implies significance at the 5% level or better

PLATE I. PAN, ROD AND TWEEZERS.

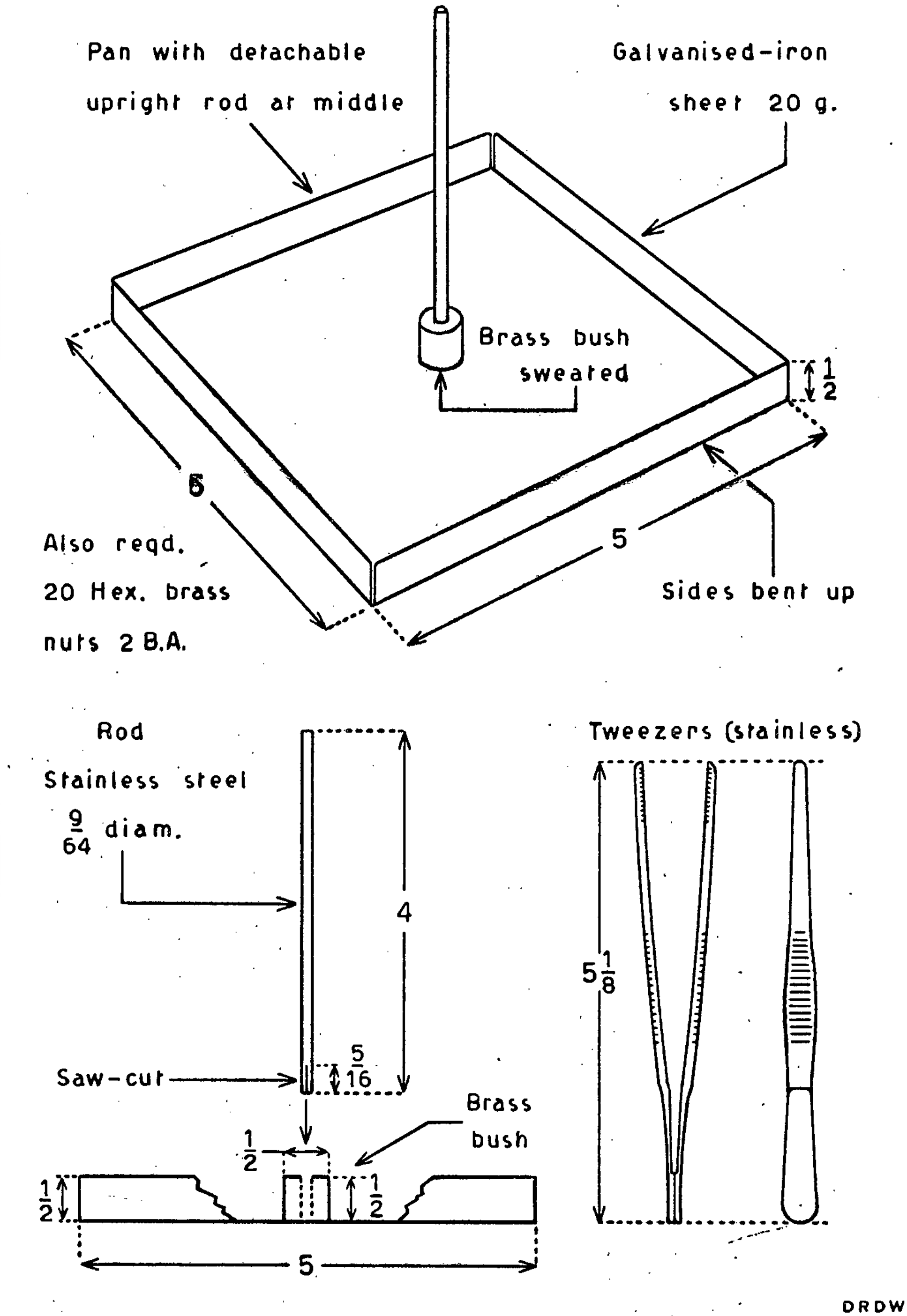


PLATE 2 STICKS AND RINGS.

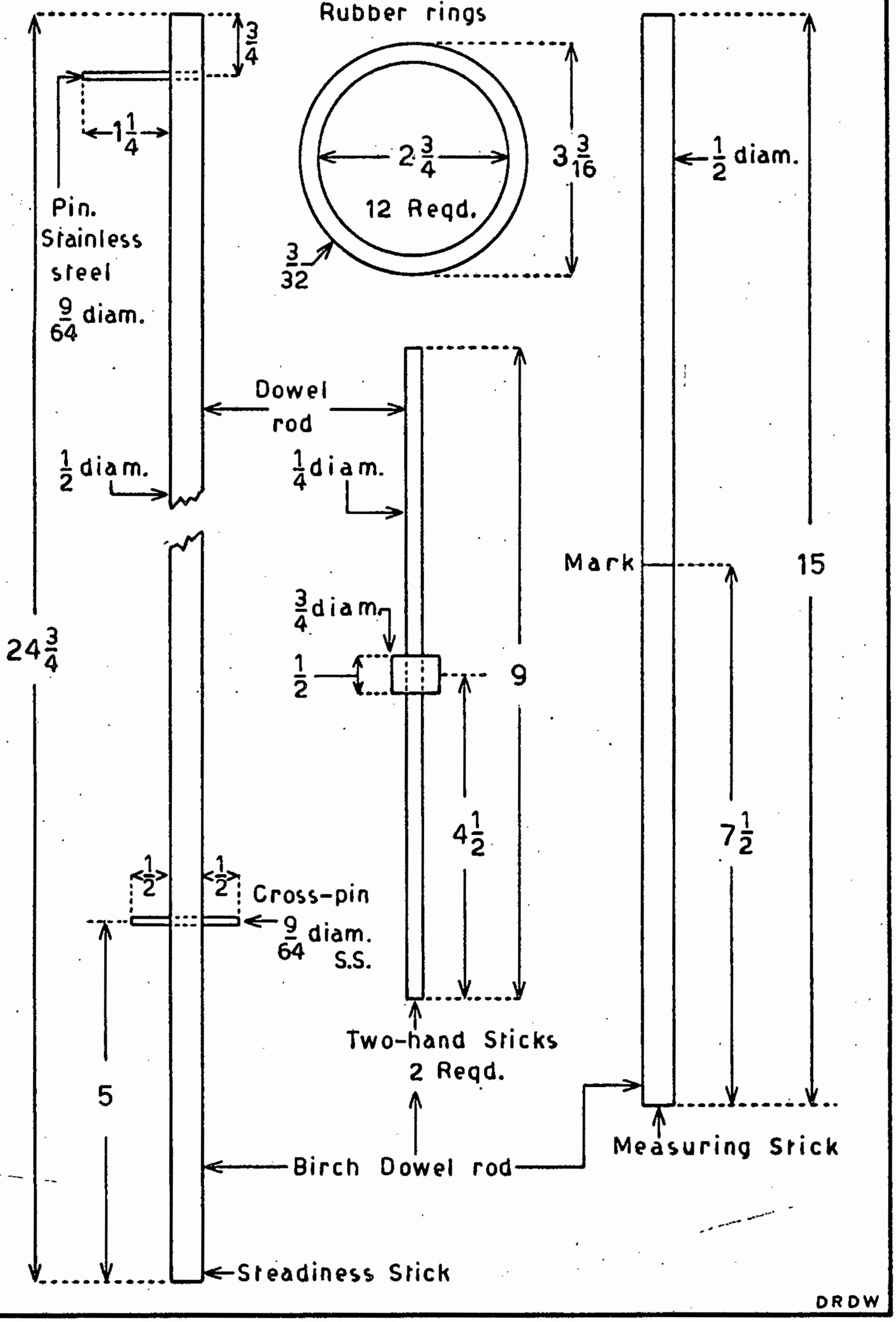


PLATE 3. MIRROR AND CASE.

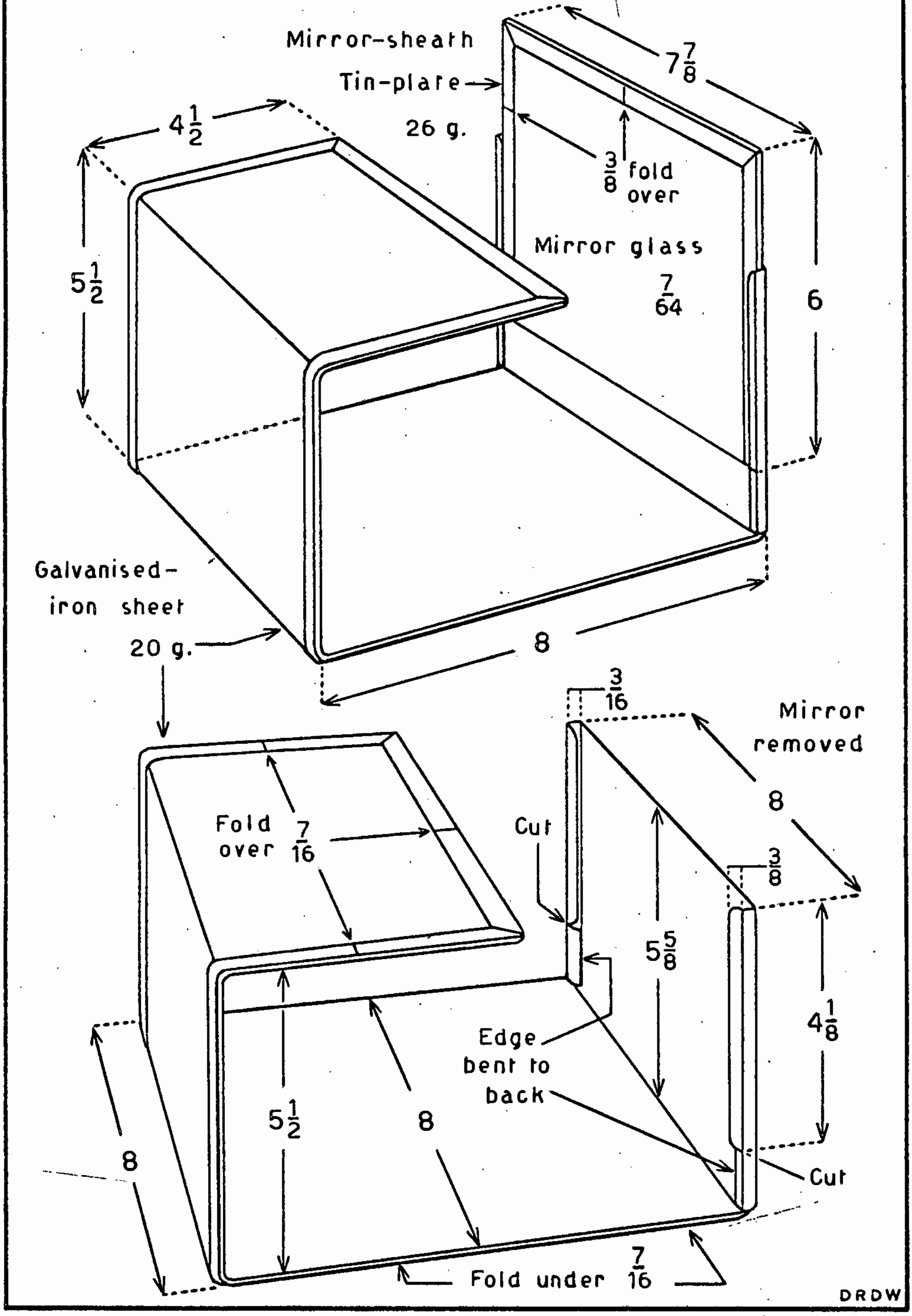


PLATE 4. CO-ORDINATION TRAY.

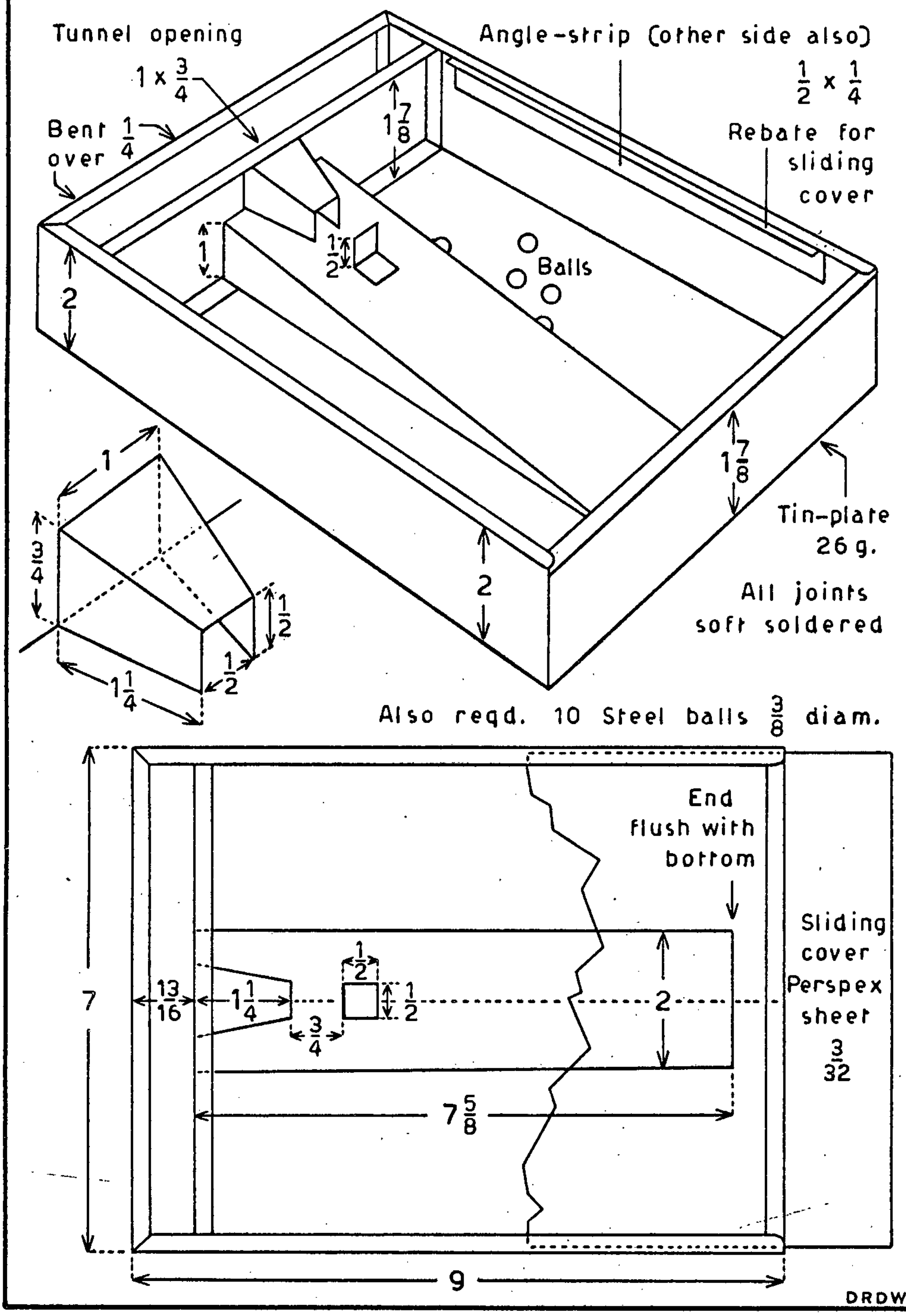


PLATE 5

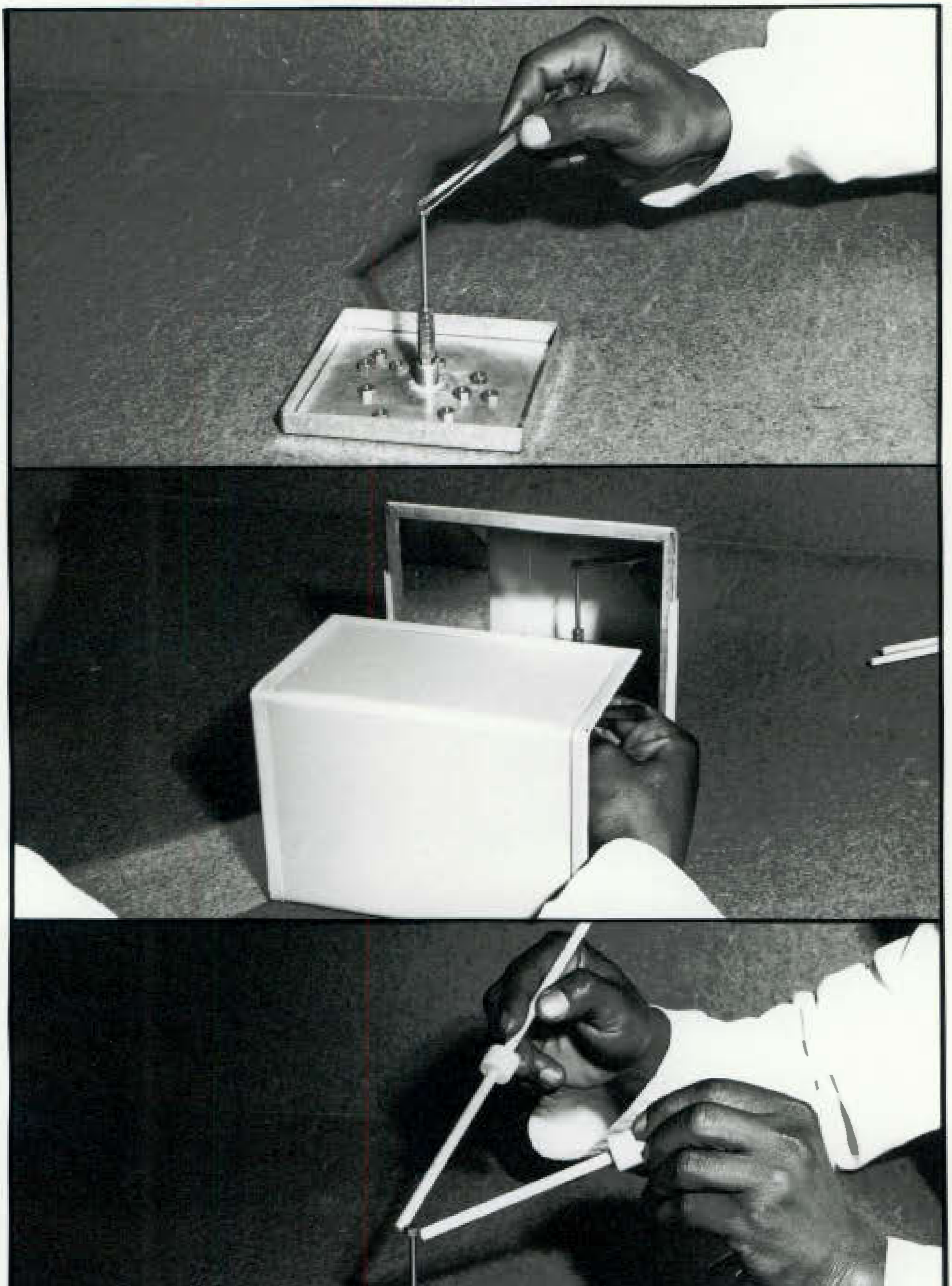
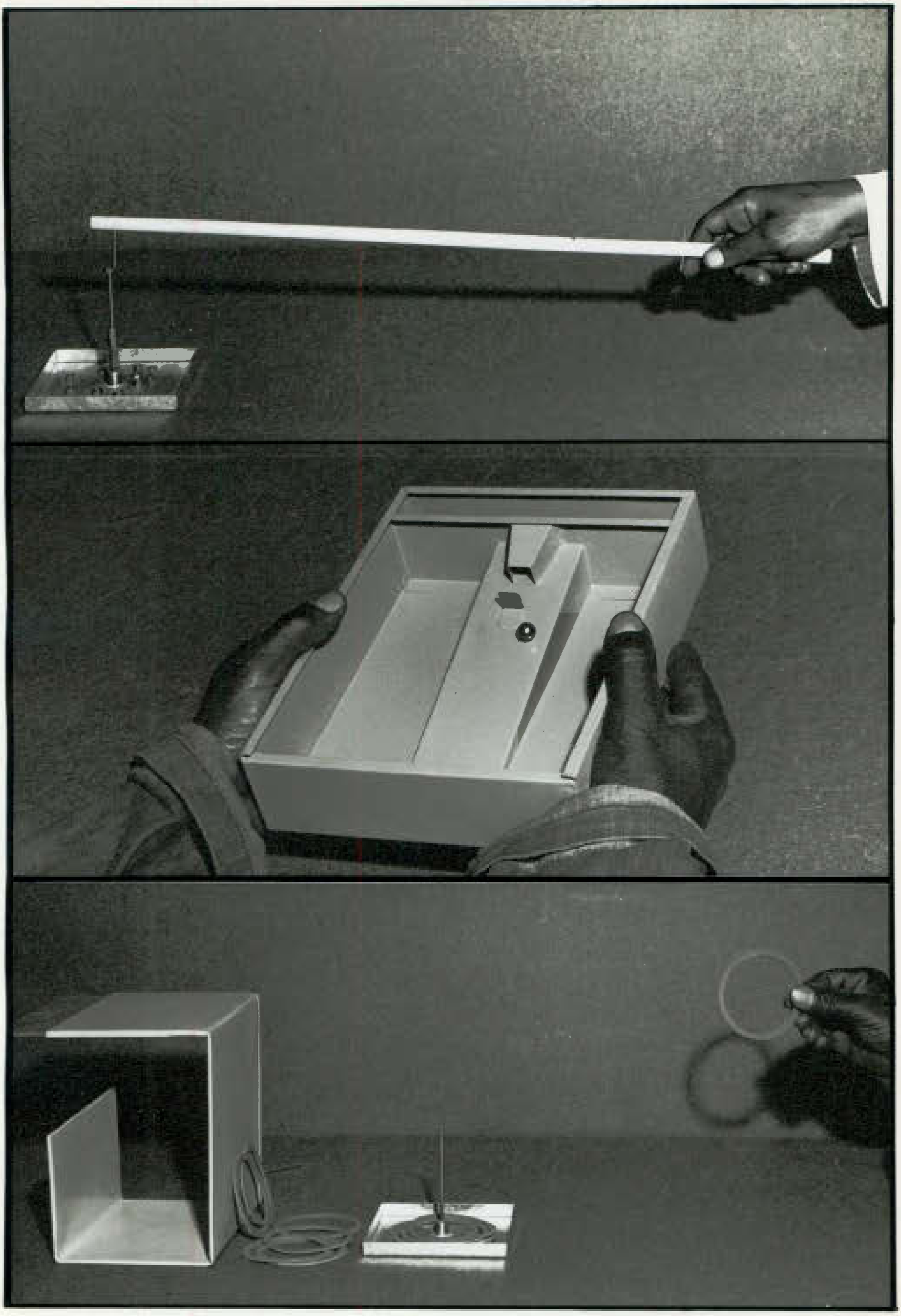


PLATE 6



ASSEMBLY AND SENSORY-MOTOR TESTS FOR AFRICAN MINE WORKERS

Largely as a result of the success achieved in the application of aptitude tests in various military and industrial spheres during the last war, the South African gold-mining industry became interested in the possibility of utilizing similar procedures for the selection and classification of their African labour, and in due course the N.I.P.R. was commissioned to undertake preliminary investigations on one of the mines.

It was soon realized that, firstly, because of the cultural heterogeneity of the subjects and the wide prevalence of illiteracy, all tests would have to be of a practical nature; secondly, most of these would have to be specially devised as few of the standard laboratory and industrial tests available were really suitable for African use; and thirdly, some tests would have to fulfil a "general-purpose" screening function and others be more specific for jobs like those of the loco- or winch-driver.

Among various apparatus tests, both "general-purpose" and "specific", which were specially devised with a view to suitability for both the subjects and the occupations, were the following: Tripod-and-weight Assembly (Model B); Object-sorting; "Three-dimensional" Peg-board; Code-signal ~~and~~ Reaction; and Dynamic Coincident Reaction.

TRIPOD-AND-WEIGHT ASSEMBLY. (MODEL B)

This unit embodies an application of one of the oldest mechanical principles known to man - the stable structure on three sloping legs capable of supporting a weight such as a cooking-pot or a lump of meat. Stability depends on the rigidity of the legs, adequate spacing of the feet, and some suitable check against the tendency of the feet to slide apart. Such a structure being common in many cultures past and present, seemed promising for a test which should preferably be as "culture-free" as possible.

The Tripod-and-weight Assembly (Model B) is a simpler and more rugged version of an earlier model which formed part of a mechanical insight battery intended for use on subjects rather more sophisticated than African mine workers.

As shown in the lower figure, recesses in a wooden block accommodate eighteen metal components and ensure a standard arrangement of presentation. The components consist of three legs six inches long made from rectangular bar (½" x ¼"), each having a hook seven-eighths from the lower extremity and a hole (3/16" diam.) at the upper; a tripod-head of 1¾" diam. x ½" stock shaped to form three equidistant forks and having a hook fixed in its centre; a cylindrical weight 2" x 1 7/16" diam. fitted with a hook of 1/8" diam. rod at the centre of one end; a ring 1¼" outside diameter (1/8" diam. rod); six links (7/8" long) of ordinary bedstead chain (wire approximately 1/16" diam.); three cheese-headed No. 2 B.A. screws seven-eighths of an inch long, and three No. 2 B.A. hexagon nuts.

PROCEDURE

The block holding the metal components is set before the subject and he is required to produce the assembled structure shown in the upper figure. Before the first trial practical instruction is given either through the medium of mime and a demonstration sample, or a movie film. The subject scores a point for every correct connection of parts made within a time limit of three minutes. After the first trial has been completed and scored, the observer dismantles what has been assembled and replaces the loose parts in their recesses in the block. The subject is then shown either a completed sample or a projected illustration of such, for twenty seconds, and proceeds with the second trial, which lasts for two minutes. After scoring and replacing the parts in the block, the completed sample (or illustration of same) is shown for ten seconds and he then does the third trial, also of two minutes duration. When the number of subjects being tested simultaneously is large, demonstration by optical projection is preferable to the use of mime and a physical sample.

OBJECT SORTING

This apparatus-test is another example of a point made in the introductory chapter to the effect that quite ordinary commercial articles can sometimes have useful applications for research purposes.

It consists of a wooden tray twelve inches square and one and three quarters deep, divided up into sixteen equal compartments; a flat oven-pan about nine inches square and one deep; and an assortment of small hardware comprising two sizes of nails, five sizes or types of wood-screws, one type of roofing bolt, three sizes or types of rivets, three sizes of washers, and two types of nuts.

Each compartment in the wooden tray is "labelled" by having an actual sample of each item fixed to its bottom. For sorting, there are five samples of each item randomly mixed in the pan.

PROCEDURE

The apparatus is set before the subject as shown in the upper figure and he must sort the items from the pan into their appropriate tray compartments, as in the lower figure. Instruction is given by actual demonstration, or a movie film, if the group is large.

The first trial lasts for two and a half minutes, and the subject's score is the number of items correctly sorted. There is no deduction for items incorrectly sorted; they are simply ignored. All items in the tray compartments are then put back randomly in the pan, and without any further instruction the subject proceeds to do the second trial, which lasts for one and a half minutes and is scored in the same way.

An improvement suggested by Dr. Hudson became standard practice on all the later batches of the apparatus. This was the replacement of the oven-pan by a wooden tray similar in external format to the one with compartments, but without compartments. At the end of a trial this "blank" tray is placed edge to edge over the other, and by turning them upside-down the items are quickly transferred from the latter to the former, where, after being shaken to mix them up well, they are ready to be sorted again. The convenience of this system, particularly when many sets of the test are being administered to a group, more than justifies the cost of the additional tray.

"THREE-DIMENSIONAL" PEG-BOARD.

Most form- and peg-board tests impose a two-dimensional task in that the subject has to consider the length and breadth of objects and recesses. But material reality is not "Flatland", however ingeniously it may be represented as such. The ubiquitous third dimension must constantly be taken into account, not only in a specialized human activity like the construction of a machine, but in the myriad activities of all organisms in relation to their physical environment.

The "Three-dimensional" Peg-board provides a test that is really a transition between conventional form- and peg-boards and actual mechanical-assembly models. As in the latter class, the three criteria of length, breadth, and height must be conformed to, but there is no catenary relationship between the various parts like that in a dynamic mechanism such as an engine or an immobile entity such as a bridge.

The apparatus (Plate 1) consists of a board (9½" x 8") with recesses, thirty special pegs, each of which is designed to fit a particular recess, and a tray for presenting the pegs to the subjects in a standard manner.

As shown in Plate 2, the board is a composite structure of two one-sixteenth brass plates with a half-inch bakelite section in between. Ten of the recesses extend all the way through; four have baffle-pins projecting up from the lower sheet, one of these being located asymmetrically; one has a transverse asymmetrical baffle-pin the end of which shows as a round dot on the edge of the bakelite near the left front corner. The small grouped recesses in which double- or triple-ended pegs fit, extend only a certain distance into the bakelite. The remaining recesses, going as far down as the lower plate, are without baffles.

Details of the brass pegs will be apparent in Plate 3. A peg with shoulders is intended for a recess open at the bottom; in a closed recess of the same size it will stand proud; whereas a peg without shoulders will fall through an open recess. A peg with a cavity will fit an unobstructed recess, but the subject will then find that he is left with a similar peg without a cavity which stands proud in the corresponding recess with the baffle. In the two recesses having asymmetrical baffles, the pegs will only fit one way, and the same holds for three of the pegs with rectangular ends and all six of the pegs with more than one end-piece, shown in the top row in Plate 3.

PROCEDURE

The upper figure in Plate 1 shows the test as it is presented to the subject, who is required to place the pegs in their proper recesses as quickly as possible and ensure that all tops are at the same level above the surface of the board. Instruction may be by actual demonstration or a movie film. The subject works for a continuous period of ten minutes and his score is the number of pegs correctly placed.

CODE-SIGNAL REACTION

A very essential branch of under-ground activity in gold-mining is the operation of winch-driven scrapers to remove the broken material from the working face and dump it at convenient localities near the end of the stope from which it can be loaded on to the coco-pans. The scraper is attached to a steel cable that is wound by the drum of a stationary winch through pulleys anchored at suitable points.

Pre-arranged code-signals sent from either the working face or the dumping station to a bell or light near the winch indicate to the operator that he must set the winch moving or stop it for special purposes, including emergencies. Promptness and precision in interpreting the signals and reacting on the controls is desirable, not only for efficient working, but also to avoid accidents.

The following apparatus was constructed to measure choice-reaction to signals given in an elementary code.

As shown in the illustration, a wooden chassis has four finger-keys mounted at one end for the observer, and three for the subject at the other. Opposite the observer's keys a flat metal frame open at one edge holds a list of the stimuli to be presented, which is protected by a sheet of celluloid. At the lower edge of the list, numbers from one to four are located opposite their respective keys. A narrow frame of similar construction holds a strip with three code-marks, each in line with one of the subject's keys; the first being one small dot and a large, the second being two small and a large, the third, three small and a large.

A vertical screen at the middle of the chassis prevents the subject from seeing a "Standard" clock calibrated in seconds and hundredths, a reset counter, two toggle-switches, and the movements of the observer in operating either the controls here or the stimulus-keys. The screen is strong and rigidly mounted in between the sides of the chassis to provide adequate support for the counter. One toggle-switch controls the main power, and the other has two settings, labelled respectively "Light" and "Buzzer". On the subject's side of the screen there is a panel-lamp in the top of the chassis.

The apparatus works off 230/250 volts A.C., but a small transformer inside steps this down to a safe six volts for the exposed key-circuits. The flex from the main supply is connected to the case at a junction-block below the front of the clock.

PROCEDURE

Test Series A consists of presentations that are entirely auditory. Having put on the "Main" switch, turned the other on to "Buzzer", and set the clock and counter to zero, the observer refers to the first item on his list, which reads "113". This means that two decided taps must be given on the code-key (number one) followed by a sustained depression of key number three, producing two short sounds and a long on the buzzer. As soon as key number three is depressed, the clock starts recording, and stops when the subject responds on the correct key for the signal given, which in this instance, would be his middle key. A reaction on any key registers as one unit of the counter, but it must be on the correct key to stop the clock.

The observer either resets the clock and counter to zero after noting the readings, or leaves the readings up if the subject's average score for say ten trials is required, and proceeds to administer the next stimulus, which on the present list, is "12". Hence, one tap on the code-key (number one) is followed by sustained depression of key number two, which produces a short and a long sound, and the correct key for the subject's reaction is that on the right. When the stimulus is "1114", as in the fourth presentation, three rhythmic taps on the code-key precede the depression of key number four, and the subject must react on his left key to stop the clock. Theobserver maintains a uniform interval of about three quarters of a second between all the components of the stimulus, which is not difficult to achieve with a little practice.

In test Series B, the procedure is the same as for A, except that the toggle-switch under the counter is set over to the position marked "Light", and the last component of each stimulus is signalled by the panel-lamp coming on instead of a long sound on the buzzer. Subjects must react as they did in Series A, the large dot in each code-mark representing the light. Ten trials are given in Series A (Auditory) and ten in Series B (Visual).

DYNAMIC COINCIDENT REACTION

Usually the winch-operator has no direct view of the scraper when it approaches the working face or the dumping station, and in either case he has to rely on a painted mark or a piece of rag on the cable. When this indicator is in line with a fixed mark he knows that the scraper has reached its the risk of the insets being touched accidentally when the electric power is on. The cowl can easily be slid out for occasional cleaning of the insets. (Plate 2, lower figure).

The reaction unit (Plates 1, 3, and 4) is a wooden case housing a counter, lamp-and-socket, control-switch, and specially constructed reaction-key. At the rear end of the case there is a junction-block for connecting the cord from the main supply, and a socket for the plug on the stimulus unit. The counter protrudes sufficiently through the top to make its reset-knob accessible, and a translucent window of ground Perspex is located directly over the forty-watt lamp. As the apparatus works entirely off 230/250 volts A.C., adequate enclosure of the components is necessary for safe operation.

Details of the special reaction-key are apparent in Plate 4. A lever of three-eighths square brass bar pivots vertically between two short uprights of one-inch angle-iron screwed to a bakelite base. At the front end of the lever, a small round insulator of bakelite is fitted, and below this, on the base, there is a rubber buffer. The rear end of the lever carries the round bakelite finger-piece, and a sloping leather-covered block of brass fixed to the base below, limits the downward traverse.

Two upright brass pillars, each made up of three separate sections bored through centrally, are mounted on the base in line with the round insulator at the front end of the lever. A pair of steel cross-pieces have their ends interposed between the sections of the pillars, and the tops of the upper sections are bridged by a bakelite cross-piece. A steel rod inside each pillar has threaded ends for nuts that clamp the structure firmly to the base. The two steel cross-pieces are drilled in the middle and function as guides for a vertical weighted plunger rounded at the lower end where it bears on the insulator at the end of the lever, and faced at the upper end with a relay-contact stud sweated on with soft solder. Another contact stud, attached to its flat spring, is supported directly over the first by having the spring clamped between the lower part of a screw-terminal and the bakelite cross-piece and bent over in a suitable bow. A screw and washer on the plunger-weight provide the connection for a light flexible "pig-tail"; the other connection being at the screw-terminal on the bakelite cross-piece. (Plate 3, lower figure). Between the top of the weight and the upper steel cross-piece, a pad of soft leather is interposed to act as a shock-absorber.

The main purpose of this special reaction-key is to ensure that contact between the points is never more than momentary. When a quick, positive reaction is made on it, the "flying" plunger rises off the bakelite insulator, makes contact with the spring-section, and then breaks contact by falling back on the insulator. No contact occurs if the key is depressed slowly, nor can any be maintained by keeping it at the "down" position shown in the right figure in Plate 4.

PROCEDURE

The two units are arranged before the subject as shown in Plate 1, and the elevation of the rail is set at the lowest graduation with the rod in the bottom hole of the fork. Steel balls, nine sixteenths of an inch in diameter, are placed one at a time by the observer at the raised end of the channel and allowed to roll down. The subject must snap down the key so as to make contact while the ball is in line with a red mark directly opposite the brass insets in the channel. If he succeeds in doing this, there is a flash of light at the window and the counter registers one unit. When he has done five trials, the reading is taken, the counter reset to zero, the elevation of the rail set at the next hole in the fork, and another similar series performed. The procedure is repeated in this way up the scale. As the slope of the channel steepens, the balls travel faster and correct reaction becomes more difficult.

APPLICATION 1.

The Object Sorting Test, Tripod-and-weight (Model B), and Three-dimensional Peg-board were administered under the supervision of Hudson (5) to African employees on a Witwatersrand gold mine. Other tests included in the battery were Cube Construction, Formboards, Screws-and-nuts Assembly, and Letter-and-numeral Sorting.

The Cube Construction test consisted of twenty-seven one-inch wooden cubes, one being plain and the rest painted red on one, two or three sides; and four rectangular wooden blocks of different sizes painted red all over. The subject's task was to arrange loose cubes to form an "all-red" replica of each block.

Formboards were of the Ferguson type, requiring the fitting of blocks of various geometrical profile into their proper recesses in boards. The test comprised four sets graded in order of difficulty.

In Screws-and-nuts Assembly, the subject had to match screws of different diameter and thread-pitch with their appropriate nuts.

Letter-and-numeral Sorting consisted of uniform metal discs stamped with letters and numerals, either singly or in combination, and a compartmented tray identical in size and construction to that used in Object Sorting, the same principle being applied of mounting an actual sample stimulus in each compartment. Further details of the test will be found in the relevant N.I.F.R. publications (1, 7).

As the tribal origins of the subjects varied considerably, standard verbal instructions could not be given, and all the tests were administered instead through the medium of mime. Later a system of demonstration by means of a silent movie film was developed.

Using Sichel's (9) minimum reliability formula, which gives a very conservative estimate, inter-trial reliability coefficents (Table 1) were computed for all the tests except Screws-and-nuts And Three-dimensional Peg-board, where only one continuous trial was given and it was impracticable to do a retest. The lowest coefficient is that for Formboards, which is, however, still acceptable.

The jobs underground in the mine had been analysed by Hudson (5) and others and finally classified into:
(1) Non-mechanical; (2) Mechanical; and (3) Supervisory. Non-mechanical work was of the unskilled labouring kind, sometimes requiring an understanding of certain simple, specific techniques, but mainly dependent on physical stamina and gross bodily co-ordination. Mechanical work involved the semi-skilled activities of assembling pipe-fittings and other simple components or operating simple power units like winches or small

locomotives. Supervision of the workers in the Non-mechanical and Mechanical groups required not only qualities of leadership but also knowledge and practical competence in both branches.

Tables 2 and 3 give the means and S.D.'s of samples in the respective categories on each of the tests, and Table 4, the significance of the differences between the means. There is a general tendency (Table 2) for the Supervisory group to score the highest and the Non-mechanical group, the lowest. As shown in Table 4, the differences between the means of the three categories on Cube Construction are not significant, but this may have resulted because the "Mechanical" sample used in this test was too small to be reliable.

Tripod-and-weight Assembly distinguishes significantly between the Non-mechanical and Mechanical categories, but not between the Mechanical and Supervisory. On all the other tests the differences between the means are significant. Referring to the last two columns in Table 4, it will be seen that when the Supervisory and Mechanical categories are combined, all the tests, including Cube Construction and Tripod-and-weight, distinguish very significantly between these and the Non-mechanical category.

A comparison of the relative mean performances of subjects from various African tribes by Gouws (4) revealed that on all the tests except Cube Construction there were significant differences. Tribes from Moçambique and the Cape Province tended to be superior to those from other South African areas and the Tropics. No significant difference was found between any of the tribes in their performance on Cube Construction. Hudson (5, p.85) proposes that "the capacity for three-dimensional thinking involved in this test is universally undeveloped among African tribes".

No significant relationship was found between the ages of the subjects, which ranged from below twenty to over forty-five years, and their performance on any of the tests.

Most of the subjects (67.3%) had no scholastic education, and of the rest, only 17% got further than Standard 1. Although a primary consideration in devising the tests had been to make them as "culture-free" as possible, it seemed likely that subjects with more educational experience and hence closer contact with European culture would have some advantage in test performance. The graphs of Biesheuvel et al (1) indicate that education definitely exercises such an influence. This was most marked in Letter-and-numeral Sorting where familiarity with the symbols greatly facilitated their identification, and also significant in Object Sorting, Formboards, and Three-dimensional Peg-board, which require an appreciation of various geometrical forms and relationships that are more characteristic of European than African culture. Education played a less significant part in Cube Construction, Screws-and-nuts, and Tripod Assembly, but the curves for these tests also showed a general upward trend. Biesheuvel (1, p.82) expresses the view that: "The tests remain primarily measures of potential ability, the level of performance being enhanced, but not determined by the rudimentary schooling which a few mine natives have received."

To investigate the possible effect of mining experience on test scores, the subjects were divided into ten groups according to their length of service and graphs plotted for the respective group means and fiduciary limits. (Biesheuvel, 1).

On Letter-and-numeral Sorting there was a slight but significant upward trend in the first three groups only, probably because it is during the earlier stages of mining experience that the illiterate men learn to identify letters and numerals, which is necessary for understanding, among other things, the implication of printed direction and warning signs.

On Screws-and-nuts there was a very significant upward trend. This was the only test in the battery having a fairly specific relationship to activity in the job situation and it is likely that the amount of preceding vocational practice at manipulating different types and sizes of screw-threads markedly affected achievement on the test.

Performance at Object Sorting, Tripod-and-weight, Cube Construction and Three-dimensional Peg-board was not influenced by the length of previous mining experience. Although, in Object Sorting, it could be expected that subjects with longer experience at working with different kinds of "hardware" would do better, this does not occur, probably because the test requires finer and more varied discriminations than are usually necessary on the job.

In a test designed to measure inherent ability, it is obviously undesirable if specific knowledge and experience play an important part in determining the subject's level of performance. Although Screws-and-nuts was found to differentiate so well between the Supervisory, Mechanical, and Non-mechanical groups, it was therefore decided not to include the scores for selection purposes, but to administer this test only as an initial "buffer" to make the subjects familiar with the general testing situation.

APPLICATION 2.

With the object of investigating validity for the selection of mechanical operatives, particularly winch-drivers, the following tests were administered to African mine workers by Mkele (6) and other members of an N.I.P.R. team: (1) Object Sorting; (2) Letter-and-numeral Sorting; (3) Cube Construction; (4) Tripod-and-weight; (5) Pursuit (Two-hand); (6) Two-hand Co-ordination (Moede type) Time Score; (7) Two-hand Co-ordination (Moede type) Error Score; (8) Dynamic Coincident Reaction; (9) Code-signal Reaction (Light); (10) Code-signal Reaction (Sound); (11) Variable Co-ordination (Time); (12) Variable Co-ordination (Error); (13) Dynamic Anticipatory Reaction.

The Pursuit (Two-hand) test was basically the same as the "Carl Heinrich" model used by the U.S.Army Air Forces in the last war. A rotating disc carried a small target-stud which the subject had to keep in contact with a pointer by turning the handles on a pair of orthogonal lead-screws. In the Two-hand Co-ordination Test (Moede type), a similarly arranged pair of lead-screws controlled a platform carrying a card with a printed track under a stationary pencil. In Dynamic Anticipatory Reaction, a small trolley drawn by a string and falling weight along rails had to be brought to rest at a given point on the rails by applying a brake to the string at a suitable stage in advance to allow for the overrun through inertia.

The reliability and validity coefficients obtained with the tests are shown in Table 5. Dynamic Anticipatory Reaction showed no relationship to the criterion, suggesting that the anticipatory kind of reaction is not important in mine winch-driving where the fairly low-speed and excessive total friction in the system greatly restricts overrunning through mechanical inertia. On the other hand, Dynamic Coincident Reaction, which has a closer behavioural affinity with the actual winch-driving task in that a reaction is made when a moving indicator is more or less directly in line with a stationary reference point, gave a coefficient of 0.45.

With the exception of Dynamic Anticipatory Reaction and the "Time" measure on Variable Co-ordination, the validity coefficients of all the tests are significant at better than the 0.001 level. Of the three Co-ordination tests, namely, variable Co-ordination, Two-hand (Moede type), and Pursuit (Two-hand), the latter showed the most substantial relation to the criterion, possibly because response to the moving stimulus involved an element of compulsion that helped to sustain attention and motivation.

On the two self-paced co-ordination tests, namely, Two-hand (Moede type) and Variable Co-ordination (combined hand and foot), accuracy of performance is more closely related to the winch-driving criterion than speed. Object Sorting, requiring the recognition of three-dimensional mechanical items has a rather higher validity for winch-driving than Letter-and-numeral Sorting, which is loaded with a literacy factor that does not enter into the criterion.

It is curious that an assembly task like Tripod-and-weight should have produced the largest coefficient. Evidently an appreciation of simple mechanical relationships, combined with adequate "g", is even more important for success in winch-driving than the ability to react precisely or co-ordinate well with the limbs.

The intercorrelations obtained between the ten more valid test measures are given in Table 6. Those larger than 0.25 surpass the 0.001 level of significance. The most substantial are: 0.63 between the two Sorting Tests; 0.61 between Code-signal Reaction to a light and to a sound; and 0.60 between Cube Construction and Tripod Assembly.

In the two Sorting Tests, the sensory-motor processes are practically identical, as the distance traversed by the hand between trays and compartments are about the same, and both the inscribed discs in the one test and the miscellaneous items of hardware in the other can be handled comfortably, there being no necessity to fumble with extremely small or flat objects affording little purchase to the fingers. The recognition of the stimulus-pieces in both tests depends on visual perception, and the main difference is that in the one set the configurations are two-dimensional, and in the other, three-dimensional.

As Code-signal reaction to either a light or a sound involves a choice-reaction series where the time taken in making the selection of the correct response is considerably greater than that required for the simple initial perception of the stimulus, the particular sense-modality employed is probably less significant than in the case of simple reaction time, where, according to Seashore (8) the usual correlation obtained between reaction to a visual and to an auditory stimulus is of the order of 0.45.

Cube Construction and Tripod Assembly probably have some common elements of "g" and the "mechanical-spatial" function recognized by Cox (2), Vernon (10) and others. This latter function could also be present in Object Sorting, which has a correlation of 0.50 with Tripod Assembly.

The most unique test appears to be Two-hand Co-ordination (Moede type), the average correlations with the other tests being very small. Particularly remarkable is the fact that Two-hand (Moede type) Time has almost no correlation with Pursuit (Two-hand), and Two-hand (Moede type) Error, a correlation of only 0.26. The essential difference between these tests is that in the Mode-Type performance is self-paced, and in the Pursuit-type, auto-

-matically paced. The stress element in the automatic pacing may well have introduced temperament factors that tended to over-run the skill factors common to both tests. There is also the likelihood that Pursuit (Two-hand) made more demands on general intelligence and mechanical-spatial comprehension, for it correlates more substantially than Two-hand (Moede type) with the Assembly tests (1,2,3 and 4). Fleishman (3) has reported the presence of a non-motor factor identified as "mechanical experience" in the similar Pursuit (Two-hand) test applied to American Air Force recruits.

Tripod Assembly shows the largest average correlation with the other tests (0.44); it also had the best validity for winch-driving (0.60). These results are probably related to the wide coverage of general intelligence, mechanical-spatial comprehension, and manipulative skill afforded by this test, particularly to an African industrial sample that was also culturally heterogeneous.

The following multiple correlations were obtained between three groups of the tests and the winch-driving criterion: Numbers 1,2,3 and 4 (known as the "Screening Battery"), 0.66; Numbers 5,6,7,8,9 and 10 (known as the "Psychomotor Battery"), 0.71; all combined, 0.74. Although the inclusion of the psychomotor tests made an appreciable increase to the predictive value of the battery, it was decided that as they involved pieces of apparatus comparatively costly to build, maintain, and administer, future selection for winch-drivers should be confined to the "Screening Battery", which consists of much simpler devices that can be reproduced fairly easily and administered as group tests.

SUMMARY AND CONCLUSIONS

(1) Five apparatus tests designed for the classification and selection of African personnel in the gold-mining industry, are described and illustrated. Three of these, namely, Tripod-and-weight Assembly, Object Sorting, and Three-dimensional Peg-board, are of the more "general-purpose" type, whereas Code-signal Reaction and Dynamic Coincident Reaction are rather specific to the occupation of winch-driving. Results from experimental applications, including reliability, validity, and relationships to other tests, are discussed.

(2) Although the tribal origins of the subjects varied considerably, and it was therefore impracticable to depend on standard verbal instructions, the tests could be administered mainly through the medium of mime or a short movie film.

(3) Tripod-and-weight Assembly (Reliability 0.83) distinguished significantly between the Non-mechanical and Mechanical Categories, but not between the Mechanical and Supervisory. Object Sorting (Reliability 0.88) and Three-dimensional Peg-board (Reliability unknown) distinguiched significantly between all three categories.

(4) On all the "general-purpose" tests except Cube Construction subjects from Moçambique and the Cape Province tended to do better than those from other South African areas and the Tropics. Differences on Cube Construction were not significant.

(5) The age of the subjects, ranging from below twenty to over forty-five years, bore no relationship to their test performance.

(6) The subject's level of education markedly influenced their performance in Letter-and-numeral Sorting, Object Sorting, Formboards, and Three-dimensional Peg-board.

(7) Length of previous mining experience had a very significant effect in Screws-in-nuts Assembly; a moderate effect in Letter-and-numeral Sorting; and none in the other tests.

(8) The most valid of the "general purpose" tests included in the winch-driver battery were Tripod-and-weight Assembly (0.60) and Object Sorting (0.50).

(9) Of the various sensory-motor tests applied to the winch-drivers, the most valid were Dynamic Coincident Reaction (0.45), and Code-signal Reaction (Light, 0.53; Sound, 0.49).

REFERENCES

1. (Biesheuvel, S.) South African Council for Scientific and Industrial Research. National Institute for Personnel Research. Aptitude tests for native labour on the Witwatersrand gold mines. Johannesburg, N.I.P.R., 1949.

2. COX, J.W. Mechanical Aptitude. Its existence, Nature and Measurement.

3. FLEISHMAN, E.A. Testing for Psychomotor abilities by means of apparatus tests. Psychol. Bull., 1953, 50, 241-262.

4. GOUWS, D.J. The influence of tribal origin on the test performance of native mine labour. Bull. Nat. Inst. Personnel Res., 1950, 2, No.2, 77-90.

5. HUDSON, W. The Occupational Classification of Africans. Ph.D. thesis. Johannesburg, Univ. of the Witwatersrand. 1953.

6. MKELE, N. Validation of Aptitude Tests for the selection of winch-drivers on the Witwatersrand gold mines. J. Nat. Inst. Personnel Res., 1953, 5, 100-109.

7. NATIONAL Institute for Personnel Research. Tests for the occupational classification and selection of industrial workers. Johannesburg, N.I.P.R. (S.A.C.S.I.R.), 1958.

8. SEASHORE, R.H. Work and Motor performance. In: Stevens, S.S., ed. Handbook of Experimental Psychology.

pp. 1341-1362. New York, John Wiley, 1958.

9. SICHEL, H.S. Note on reliability of combination of subtests, tests, or criteria. Bull. Nat. Inst. Personnel Res., 1950, 2, No.2, 57-60.

10. VERNON, P.E. The structure of practical abilities. Occup. Psychol., 1949, 13, No.2, 81-96.

TABLE 1

Inter-trial Reliability Coefficients

Tests	r	No.of Cases	S.E.
Letter-and-Numeral Sorting	0.911	365	0.009
Object Sorting	0.881	347	0.012
Tripod-and-Weight	0.830	363	0.016
Cube Construction	0.785	365	0.020
Formboards	0.704	360	0.027

TABLE 2

Mean Scores for the Major Job Categories

Tests	Non-Mech.		Mech.		Supervisory	
	N	Mean	N	Mean	N	Mean
Letter-and-Numeral Sorting	595	44.0	258	53.5	257	78.0
Object Sorting	563	65.0	248	76.2	257	85.0
Tripod-and Weight	193	25.8	88	35.3	257	38.1
Cube Construction	60	18.5	11	23.8	257	26.7
Formboards	596	21.2	260	24.1	556	25.9
Screws-and-Nuts	597	6.7	259	8.7	558	10.7
Three-Dimen. Peg-board	536	12.3	235	14.9	257	16.9

TABLE 3

Standard Deviations of the Means

in Table 2

Tests	S.D.'s		
	Non-Mech.	Mech.	Supervis.
Letter-and-Numeral Sort.	28.6	32.0	32.7
Object Sorting	34.4	37.3	31.5
Tripod-and-Weight	15.3	14.2	13.3
Cube Construction	11.4	9.7	12.2
Formboards	9.1	9.4	9.3
Screws-and-Nuts	3.0	3.7	3.5
Three-Dimen. Peg-board	7.2	7.4	7.2

TABLE 4

Critical Ratios of the differences between the means of the Major Job Categories.

Tests	Non-Mech. and Mech.		Mech. and Supervisory		Non-Mech. and Superv.+ Mech.	
	C.R.	P	C.R.	P	C.R.	P
Letter-and-Numeral Sort.	4.1	<.001	8.6	<.001	11.8	<.001
Object Sorting	4.0	<.001	2.9	.004	7.5	<.001
Tripod-and-Weight	5.0	<.001	1.6	.110	8.8	<.001
Cube Construction	1.6	.110	1.0	.317	4.9	<.001
Formboards	4.2	<.001	2.6	.009	8.3	<.001
Screws-and-Nuts	7.5	<.001	7.4	<.001	19.3	<.001
Three-Dimen. Peg-board	4.5	<.001	3.0	.003	8.0	<.001

TABLE 5

Reliability and Validity (N=150)

Test	Reliability	Validity
1. Object-Sorting	.88	.50
2. Letter-and-Numeral Sorting	.91	.44
3. Cube Construction	.79	.49
4. Tripod-and-weight	.83	.60
5. Pursuit (Two-hand)	.89	.42
6. Two-Hand Co-ord.(Moede-type) Time	.85	.31
7. Two-Hand Co-ord.(Moede-type) Error	.93	.38
8. Dynamic Coincident Reaction	.68	.45
9. Code-Signal Reaction (Light)	.89	.53
10. Code-Signal Reaction (Sound)	.94	.49
11. Variable Co-ordination (Time)	.93	.11
12. Variable Co-ordination (Error)	.93	.30
13. Dynamic Anticipatory Reaction	.71	.08

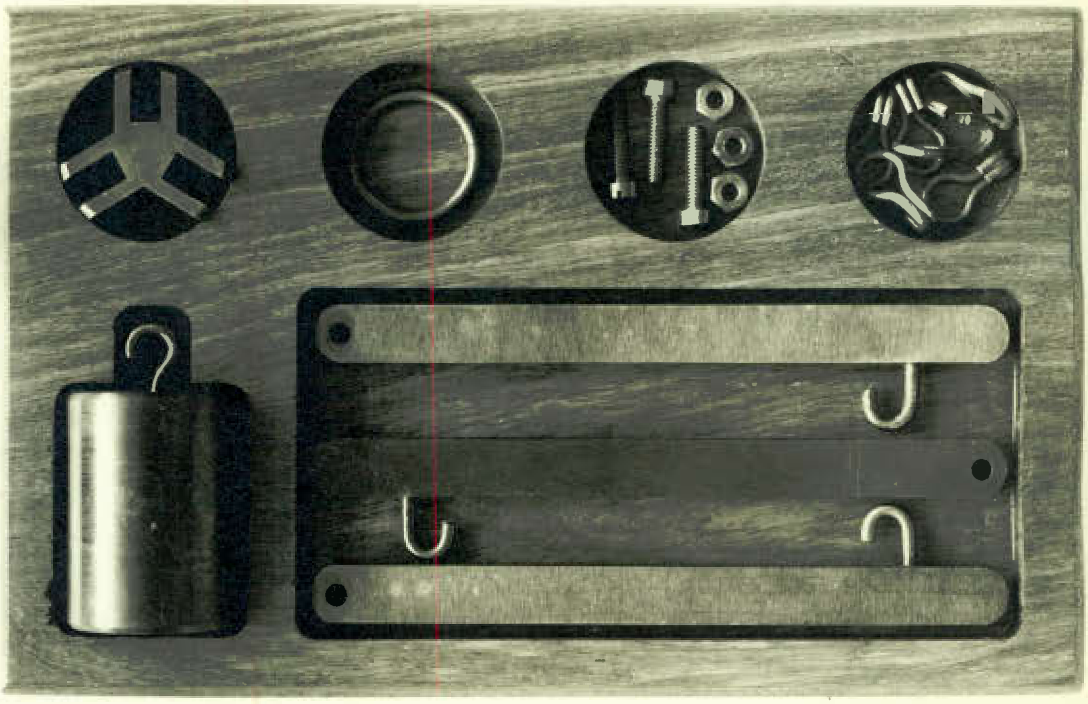
TABLE 6

Intercorrelations between tests -

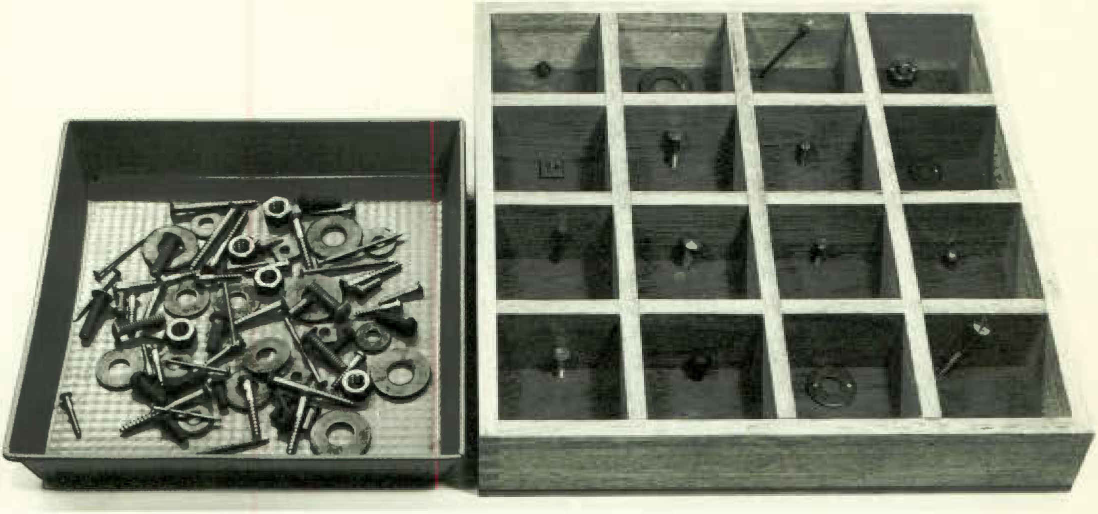
(Decimal points omitted) (N=150)

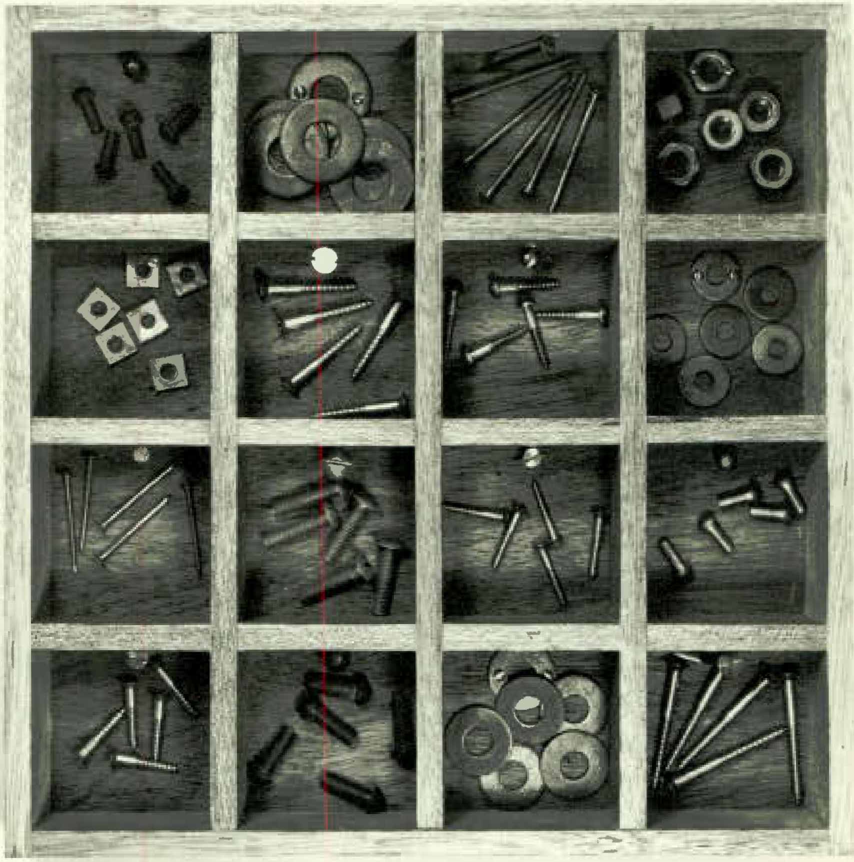
Test	1	2	3	4	5	6	7	8	9	10
1 Object-sort.										
2 Letter-and-Numeral Sort.	63									
3 Cube-construction	33	42								
4 Tripod-and-weight	50	46	60							
5 Pursuit (Two-hand)	40	38	36	46						
6 Two-hand (Moede) Time	14	16	21	27	09					
7 Two-hand (Moede) Error	31	24	19	31	26	13				
8 Dyn.Coin. Reaction	25	26	26	43	22	21	15			
9 Code-signal (light)	37	44	39	44	26	25	23	44		
10 Code-signal (Sound)	31	43	41	46	26	09	26	27	61	
Mean r	36	38	35	44	30	17	23	28	38	34

TRIPOD-AND-WEIGHT ASSEMBLY. MODEL B.

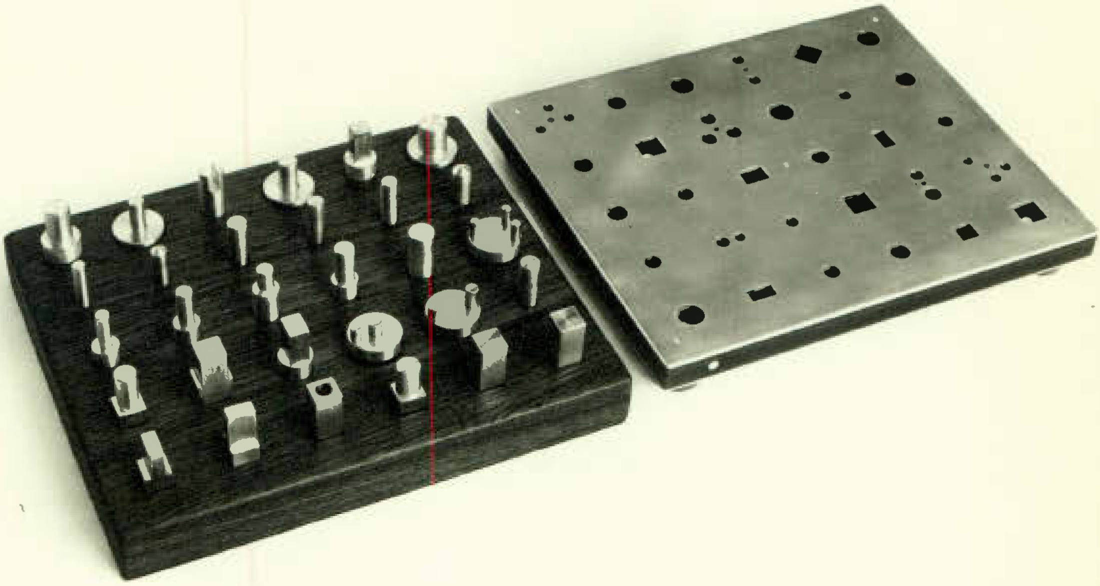


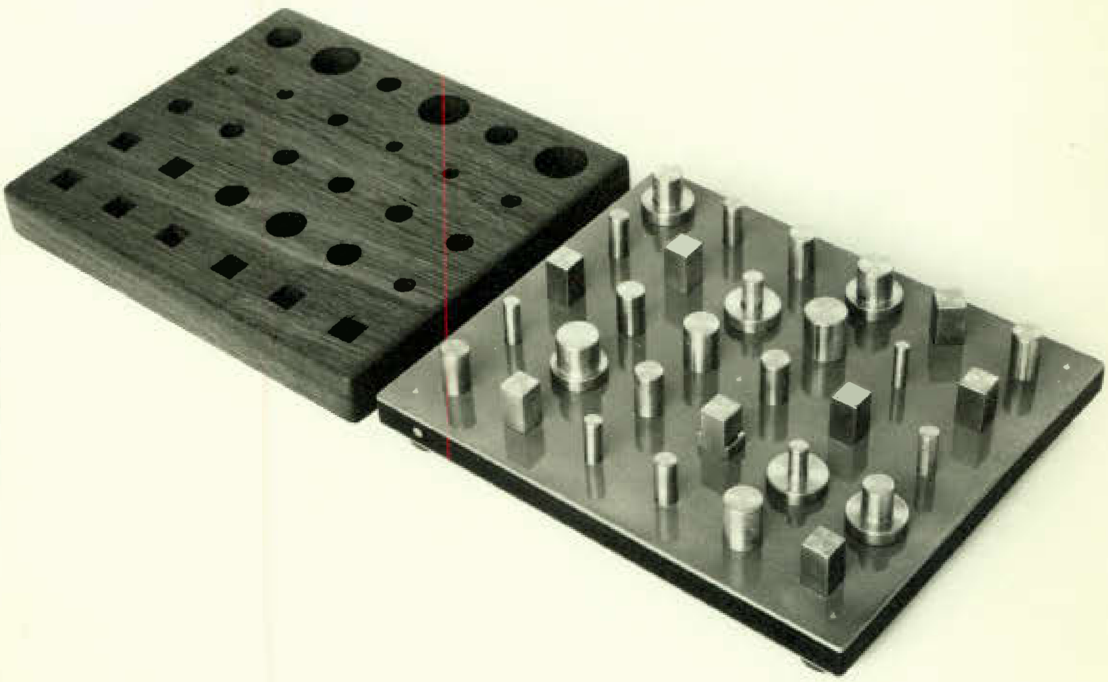
OBJECT - SORTING TEST.



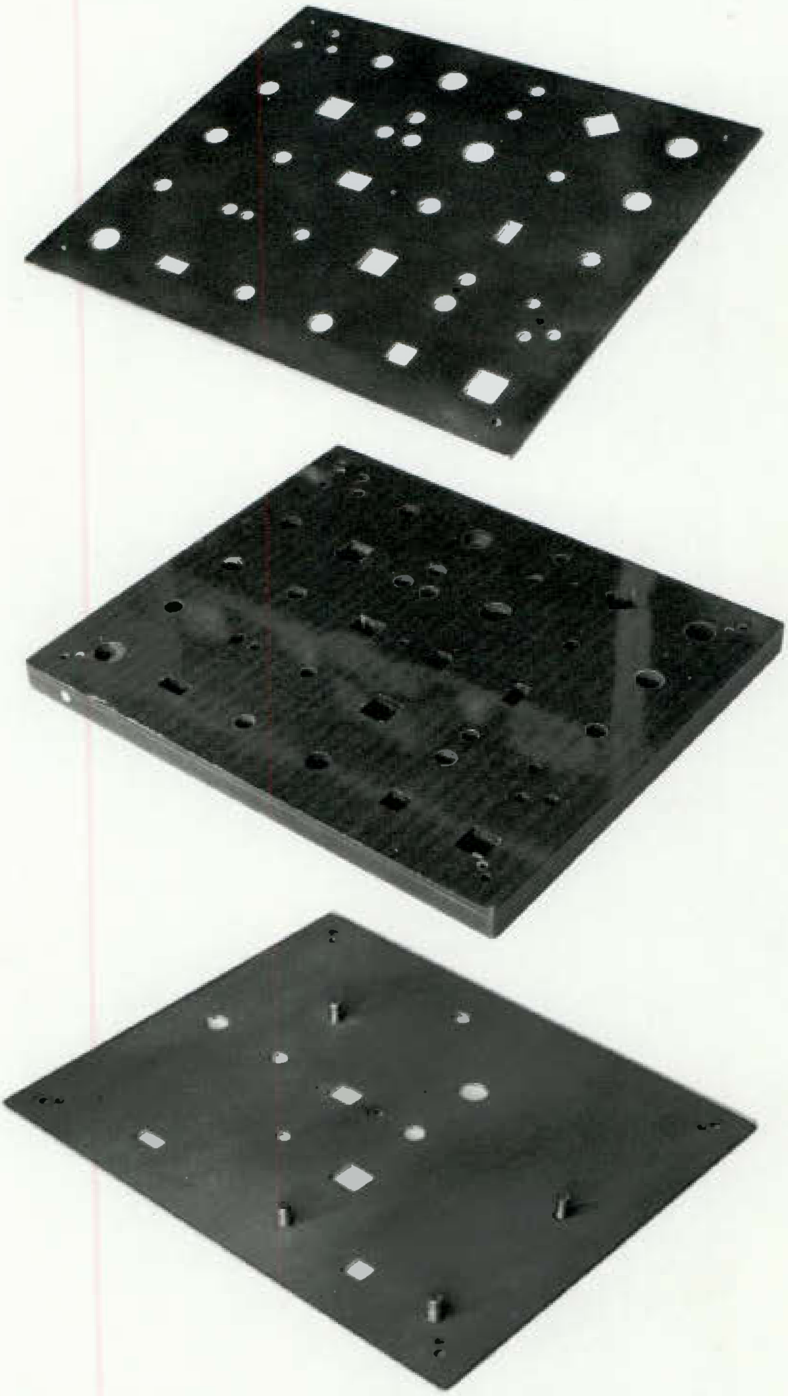


"THREE - DIMENSIONAL" PEG - BOARD. PLATE 1.



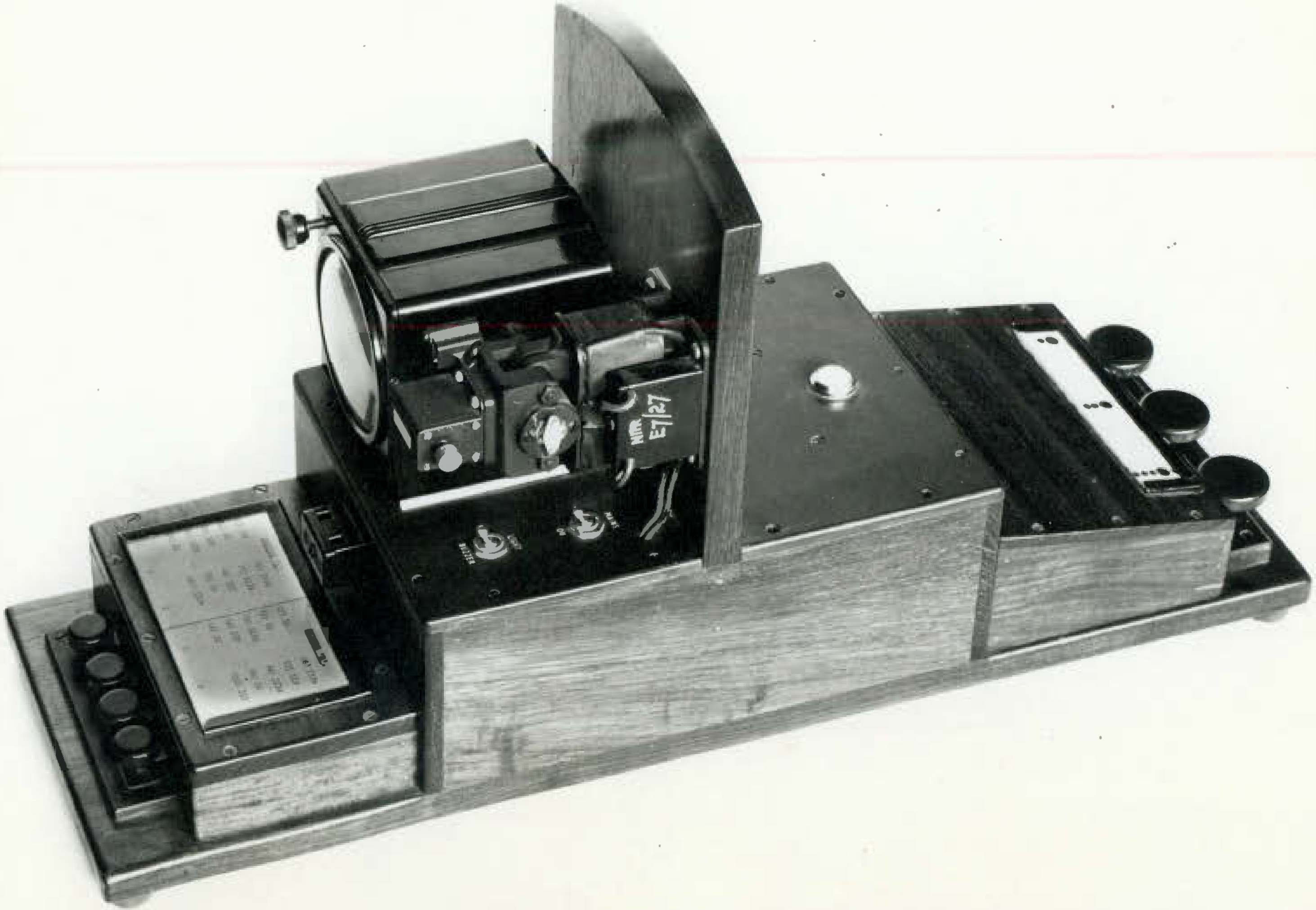


"THREE-DIMENSIONAL" PEG-BOARD. PLATE 2.

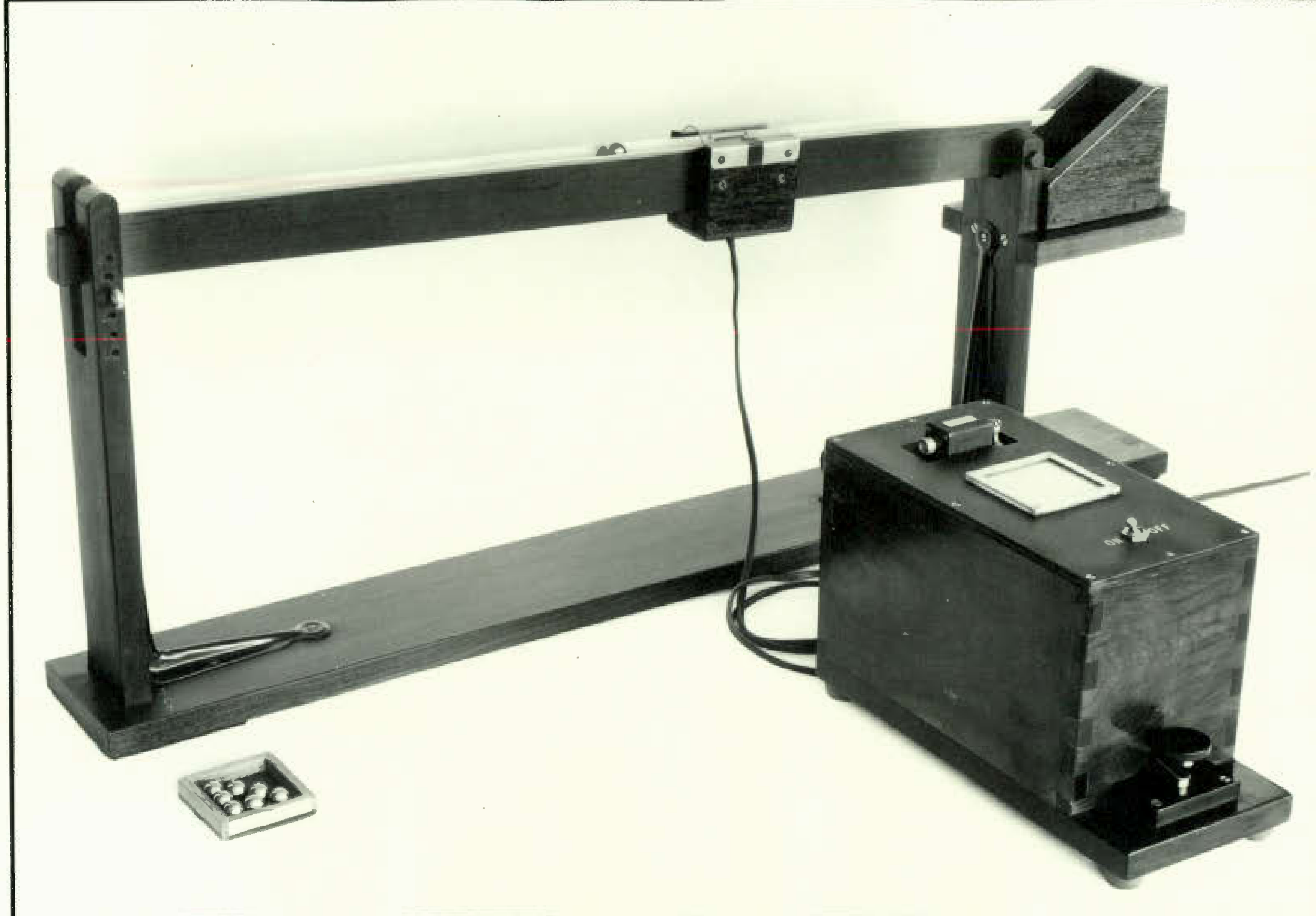


"THREE - DIMENSIONAL" PEG - BOARD. PLATE 3.

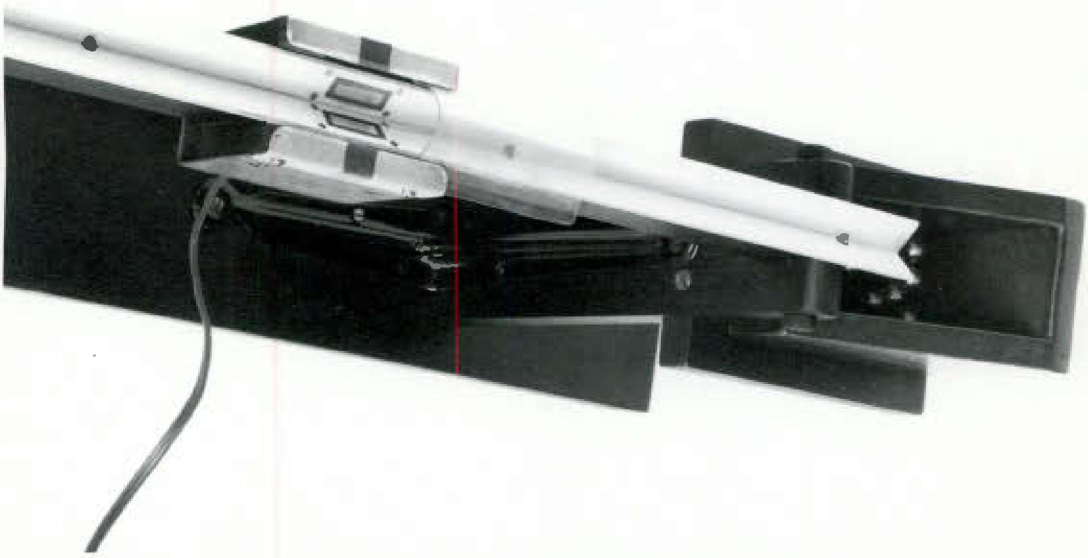
CODE-SIGNAL REACTION.



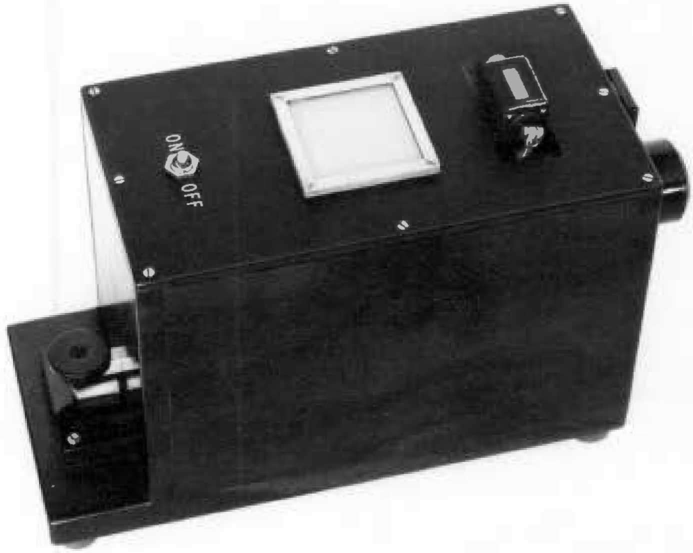
DYNAMIC COINCIDENT REACTION. PLATE 1.

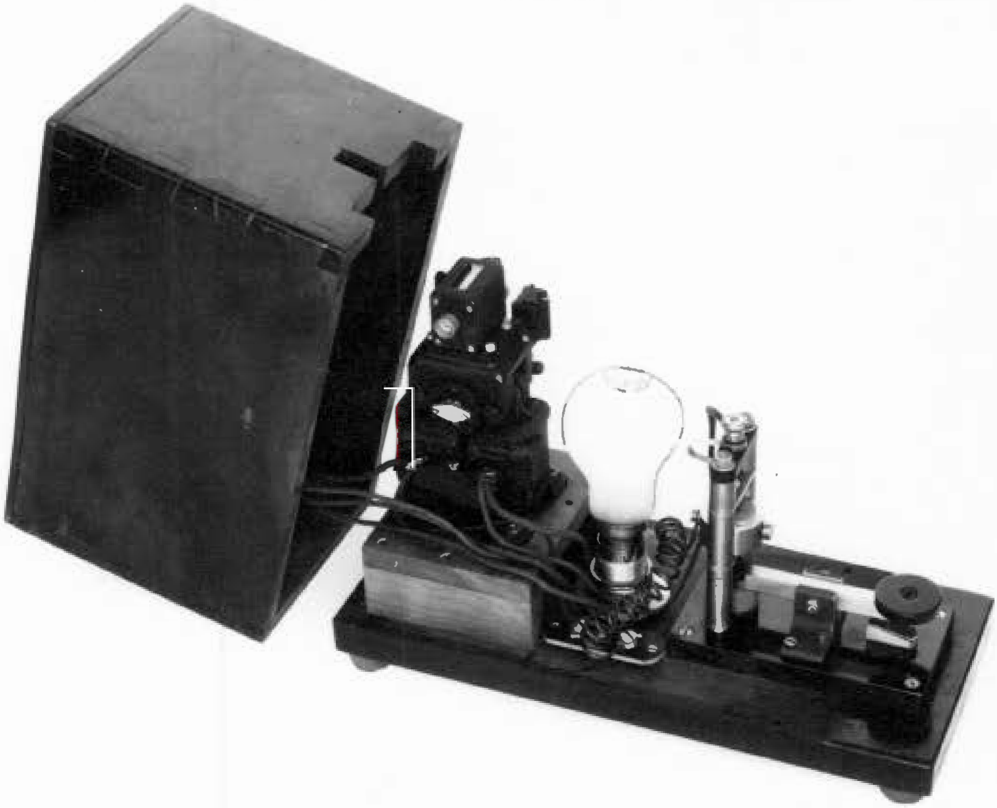


DYNAMIC COINCIDENT REACTION. PLATE 2.



DYNAMIC COINCIDENT REACTION. PLATE 3.





DYNAMIC COINCIDENT REACTION. PLATE 4.

SECTION II

Summary

1. Technical and research details are furnished of two tests of fine Implement manipulation; six tests of moderately fine Manipulation, of which four involve the use of implements; three Assembly tests, and two Reaction tests.

2. Inter-trial correlations show that twelve of these thirteen tests have useful reliability. It was not possible to assess the reliability of the "Three-dimensional" Peg-board, because it was given as one continuous task. Chopsticks, which was administered more than once, also showed good Test-Re-test reliability. In this test repetitions tended to improve both internal consistence and Test-Re-test reliability, but the introduction of a monetary incentive, while beneficial to all-round achievement, reduced internal consistence, possibly because subjects were sometimes tempted to "take chances".

3. Achievement in Chopsticks and Tweezer-Nozzle correlated respectively, 0.59 and 0.51 with achievement in a factory operation of "Nutting" (with incentive); without incentive, these correlations were 0.51 and 0.38.

4. Results from Chopsticks and the Nutting operation over a 21 day period showed that practice and training did not even out individual differences in initial ability. Those subjects who were good initially, tended to keep ahead.

5. Chopsticks showed small positive relationships with Secondary Function, Emotional Stability, Visual Selective Reaction and Visual Perception Rate; and small negative relationships with Heart and Respiration rate. The most significant result was a correlation of -0.36 between Chopsticks achievement and Pyruvic Acid level after physical work, probably attributable to biochemical factors influencing behaviour. Chopsticks has proved useful as a situation for observing temperamental manifestations.

6. On Tweezer-Nozzle, where it was possible to compare the respective performance of Africans and Europeans, no significant differences were found between the groups.

7. Tweezer-Nozzle correlates most significantly with Chopsticks (0.48) and Tripod Assembly (0.44), where the patterns of manipulation also require rather fine muscular control.

8. On Tweezer-Nozzle, good performers tend to estimate their achievement more realistically than inferior performers. Most subjects appear to have a more or less exaggerated idea of their ability and are inclined to over-estimate what they can actually achieve. The over-estimators tend to be further out in their estimations, than the under-estimators.

9. Although mild to moderate alcoholic intoxication caused a significant deterioration of performance in Visual Discrimination, Tapping Speed, Card Sorting and Two-hand Co-ordination (Accuracy) it had no significant effect on performance in Tweezer-Nozzle.

10. While Chopsticks and Tweezer-Nozzle both lend themselves to group administration, they are somewhat costly to reproduce in quantity. The six "simple" skill tests, namely Tweezer-Nuts, Tweezer-Mirror, Two-hand Sticks, Steadiness Stick, Tray Co-ordination, and Ring Throwing, which provide quite a wide variety of manipulative patterns, are also suitable for group administration and have the advantage of being fairly cheap to construct.

11. The fact that the six latter showed no validity for selecting operatives in a porcelain factory, could have been partly due to the lack of specific and reliable criteria. Tweezer-Nuts, Two-hand Sticks, and Steadiness Stick have since been applied by a large industrial organization to operatives in a Cashew-nut factory in Mocambique, and more definite validation results should be obtainable because the daily productivity of operatives in this factory will provide a specific and accurate criterion.

12. In Tweezer-Mirror, where the task was complicated by perceptual motor conflict, the more intelligent subjects performed better. These skill tests tend to be highly specific with a mean intercorrelation of only 0.15. Among the more intelligent subjects the correlation between mental ability (measured by a Form Series test) and the sensory-motor tests is very small (0.12). Among the less intelligent it is slightly but significantly larger (0.24). In certain of the tests there were differences in achievement attributable to differences in sex and age.

13. Of the three Assembly tests administered to mine-workers, Object Sorting and "Three-dimensional" Peg-board distinguish well between the Non-mechanical, Mechanical, and Supervisory categories; and Tripod-and-weight, between the Non-mechanical and Mechanical, but not between the Mechanical and Supervisory.

14. While no significant tribal differences were detectable with the six skill tests applied at the porcelain factory, on most of the "general-purpose" tests applied to the mine-workers there were significant differences, subjects from Mocambique and the Cape Province tending to be superior.

15. Previous formal education did not benefit the "porcelain factory" groups in their performance at the six "simple" skill tests, but it was advantageous to the mine-workers in those Assembly tests where the main factor was the ability to relate forms, namely, Sorting, Formboards and "Three-dimensional" Peg-board.

16. In none of the mine-workers' tests did the age of the subjects bear any relation to their test performance. But in the "simple" skill tests applied to the factory-workers, very young subjects (mainly females) fared the worst at Tweezer-Nuts, the initial test. At Tray Co-ordination both the youngest and the eldest scored less than an intermediate group.

17. Length of previous mining experience had a very significant effect in a test of Screws-and-nuts Assembly, a moderate effect in a test of Letter-and-numeral Sorting, but none in any other tests.

18. The most valid of the "general-purpose" tests included in the winch-driver battery were Tripod-and-weight (0.60) and Object Sorting (0.50). Of the various Sensory-motor tests applied to the winch-drivers, the most valid were Dynamic Coincident Reaction (0.45) and Code-signal Reaction (Light, 0.53; Sound, 0.49).

19. All the tests are suitable for use on polyglot groups because most of the requirements can be communicated to the subjects by demonstration through mime or a movie film.

SECTION III

AN APPRAISAL OF SENSORY-MOTOR ACTIVITIES

Introduction

Partly on philosophical and empirical grounds, and partly as a matter of convenience in research, it has been found useful by many to accept a rough classification of animal behaviour into the following five processes: (1) Sensory, concerned with the reception of physical stimulation from the environment or from within the physical structure of the organism itself. (2) Perceptual, in which the sensory data is interpreted; (3) Motor reactions of muscles and glands in response to the perceptual process, or reflexively, in direct response to the sensory stimulation; (4) Emotional, involving effective or conative sets, usually accompanied by reflex muscular and glandular activity, particularly in response to strong elementary drives such as self-preservation, mating, and aggression. (5) Intellectual, which seems to involve a complex amplification of the perceptual in that simple and rather isolated units of interpretation become associated and systematized. This process is the particular speciality of <u>homo sapiens</u>, and manifestations of it by less voluble and conceited animals like the chimpanzee, cat, dolphin, horse or dog, are usually regarded by man as not strictly "genuine".

The designation "motor" implies movement. In biological contexts, movement is something more than the straightforward result of applied mechanical force, for even at the reflex level it is often preceded by a modicum of elementary coding dependent on sensory data. In referring to the movements of animate organisms it has therefore become conventional to employ a more descriptive composite term such as <u>sensory-motor</u>, <u>perceptual-motor</u>, or <u>psychomotor</u>. The detailed investigation of the physical mechanism of cells, organs, nerves and muscles that produce such movements, falls within the provinces of physiology, anatomy and biochemistry. Psychologists, educationists, and employers of personnel are mainly concerned with the behavioural aspects of movements in relation to external things, and generally, they are more interested in voluntary than involuntary activity.

Diverse Activities

Human beings, like other animals, spend much of their lives in making muscular movements towards, away from, among and with, a variety of things. There is continual physical reaction to sensory data from a material environment, either in prusuit of a livelihood or for recreation. The most fundamental kinds of directed sensory-motor activity are those of a gross nature like walking or running, jumping, swimming, climbing, and the carrying or dragging of objects, where the main requisites are strength or speed, or a combination of both. In these fields, man is markedly inferior to the brute creation, and despite the powerful incentives provided by Olympic gold medals and consequent national deification, he cannot outrun the hare, jump further than the kangaroo, outswim the shark, climb with greater agility than the baboon, or move heavier weights than the horse. In one kind of gross activity he has a certain advantage over most animals, his hands enable him to grip an implement such as a spade or lever, but even in this respect he is physically inferior to the anthropoid ape with four prehensile members, and the elephant with a single big one.

On the score of "quantity" in sensory-motor potentiality, man is therefore a puny creature, but Nature has seen fit to compensate him well with superior endowments as regards "quality". He excels in those activities where spatial precision is important. These may also be of a gross kind, necessitating considerable strength, like moving a piano to a definite position, felling a tree with an axe, making a long place-kick in football and hitting sixes at cricket; or of a fine kind where the movements are usually of rather restricted scope, requiring little physical strength, and the greater bulk of the body is more or less static and serves as a buttress that facilitates the maintenance of balance and optimum posture. Examples of the latter are, the barber shaving either his own chin or a customer's; the dentist drilling and filling a tooth; the little girl practising her scales on the piano; her brother doing fret-work on a small treadle machine; mother darning socks; and father at his writing table struggling to reassemble the domestic alarm-clock which he was unwise enough to take apart and clean.

Among the lower creatures, spatial precision relates mainly to the judging of distance that has to be traversed quickly, in order to avoid obstacles or seize prey, but in the case of certain species it is also notable in the construction of dwellings and devices for securing or storing food. The web of the garden-spider, the den of the trap-door spider, the honeycomb of the bee, the nest of the weaver-bird, the perfect funnel-shaped pit in the sand made by the tiny, ferocious "ant-lion" to trap the unwary ant, the dams of the beaver - are remarkable for both precision of execution and efficiency in use. Such ingenious activities on the part of organisms that are ostencibly devoid of volition, we ascribe vaguely to "instinct", but until biological science is able to give some positive definition of "instinct", we cannot lightly dismiss the conclusion reached long ago by Lord Avebury, that the mental powers of certain lower creatures "differ from those of man not so much in kind as in degree."

Although it is useful to postulate a broad distinction between gross activities of the athletic type and those requiring the exercise of finer muscular co-ordinations, there are at least six principal aspects to all voluntary activity, namely, the quantitative ones of strength, speed and amplitude of movement, and the qualitative ones of accuracy, action and pacing. In "single-action" activity, separate unitary movements are made, like pressing a key whenever a bell rings, striking a cricket-ball, or using a stop-watch; in "serial action" a more or less connected sequence of similar unitary movements is made, as in sending messages by the telegraph-key, typing and knitting. Sometimes an activity may include both. For example, in playing the organ, serial-action in pressing the manual keys and the pedals is interspersed with single-action in operating the stops. In "continuous action", the movements tend to be smooth and rather sustained, as in tracing or drawing, ironing a garment, or pushing a cart. Rhythm of some sort can play a part in all kinds of action.

As regards pacing, there is the distinction between "self-paced" activities, where there is no definite standard of speed to be maintained, and although it may be necessary to move "as quickly as possible", the actual physical implication of this remains a matter of personal judgment; and "pursuit" activities where there is some overt external compulsion to move at a certain rate, which may well be rather faster than is comfortable. Examples of the first kind are, walking alone to work (briskly, if you are behind time), writing a fair copy of a letter, and sewing a piece of cloth by hand; and of the second, walking beside a long-legged friend who is in a hurry, writing notes in a lecture (particularly if the speaker proceeds rather fast), and sewing with a motor-driven machine.

Much human sensory-motor activity is concerned with either the production of material things that may contribute to physical and mental well-being, or the use of such things. To make up for his own poor quantitative endowments, man is forced to rely extensively on mechanical accessories, and has become a tool-making and a tool-using animal who constructs and applies a multiplicity of instrumental aids for securing food, eating food, obtaining protection from the elements, bringing forth and rearing young, destroying his enemies, boosting his ego, journeying far and near, and amusing himself generally.

Observation and Speculation

It is possible, that long before the science of differential psychology came into being, certain basic facts about ability in various sensory-motor activities were quite well known. Firstly, at a given activity, whether it was shaping a flint arrow-head, shooting with the completed arrow, chasing a wounded antelope, processing its skin into a garment, or carving an effigy of the animal on the cave wall - some practitioners were better than others. Secondly, a few of them were superlatively good; a few, really poor; and the majority middling. Thirdly, it was rare for an individual to excel at many widely different activities. Fourthly, practice increased proficiency. For thousands of years there was probably little speculation about such rather obvious facts, which were just accepted in the light of casual observation and common-sense. Things might, or might not have been what they seemed, but life was real, life was earnest and mere theorizing seldom produced tangible results. When not actually engaged in hunting, fashioning weapons, garments and receptacles, preparing food, eating, fighting or procreating, it was preferable to sleep and invigorate the tired frame for the resumption of these active pursuits on the morrow.

However, <u>tempora mutantur, et nos mutamur in illis</u>, and with the change in times and outlook speculation of a superior kind, that often soared out of sight above ordinary observation and common-sense, became another of man's diversions: <u>cogito ergo sum</u>, and so forth. In this absorbing new pastime, which might be called the philosophical precursor of inductive psychology, the player directed his efforts towards the establishment of general laws and principles by sitting in a comfortable armchair and thinking about numerous individual specimens, both real and imaginary. Only much later did it become fashionable to get up occasionally and submit some of the real specimens to detailed and critical examination. When the former "philosophic" approach was more in vogue than the latter "empiric", a grouping of abilities was postulated that tended to be very broad, and produced a compact, plausible picture which subsequent research has shown to be, in many respects, an inaccurate portrayal of reality. For example, the eighteenth century faculties of memory, will-power, attention or concentration etc., were found to be little more than theoretical "blanket" concepts, because experiments demonstrated that individual ability in any one of them varied greatly with the nature of the subject matter.

Much of the early work in experimental psychology comprised the intensive testing of rather small samples to establish how well the "average person" could perform various mental, perceptual and sensory-motor acts; among the latter, the measurement of simple reaction-time to several kinds of stimuli being particularly popular.

Individual differences (the "personal equation") in the performance of various tasks, had, of course, been observed, but when it was generally appreciated that somewhat larger test samples were more representative of humanity as a whole, and these were used, the significance of individual differences became accentuated and the main interest shifted to the question of the interrelation between abilities.

Sensory-motor and Mental Abilities

During the first quarter or so of the present century, there was considerable interest in the relation between "mental" and "motor" performance. Research on mental tests had shown pretty conclusively that a powerful central tendency pervaded all facets of conscious cognition, and there was adequate justification for postulating a general factor of intelligence. Largely on the strength of this, it was tacitly assumed by some investigators that a general factor of "motor" ability also existed, and in the literature of the period, titles, discussions and conclusions frequently allude to "motor" ability, capacity, power, or skill etc. in the singular.

Correlation between performance on certain mental and sensory-motor tests produced a wide diversity of results. Some investigators concluded that the relationship was approximately zero, others that it was positive, and a few, that it was actually negative. Subsequent research has tended to corroborate a dictum of Thorndike's which is aptly quoted by Seashore (59): "In general, desirable human characteristics are slightly, but positively, correlated." Seashore (59) interprets "slightly" as implying a correlation of the order of 0.15 to 0.25.

Table 1 shows a fairly typical set of intercorrelations obtained with a battery of mental "pencil-and-paper" tests on a sample of male matriculants. All coefficients, with the exception of that between Arithmetic and Aviation Interests, are significant at better than the 5% level, and the mean coefficient for the whole table is 0.40. The pattern suggests that there is, firstly, a strong general factor pervading all the tests and most concentrated in Mental Alertness and Progressive Matrices; secondly, a more specific reasoning factor in Mathematics and Arithmetic; and thirdly, a "knowledge" factor in the two Technical and Scientific tests and the Aviation Interest Inventory.

In Table 2 are the intercorrelations obtained between measures on four sensory-motor tests administered to the same sample. The three self-paced tests, namely, Steadiness (15), Hand-foot Reaction (14) and Two-hand Co-ordination (Moede-type), afforded measures of both speed and accuracy, but the Air-pilots arm-leg Co-ordinator (A.L.C.), which is a pursuit test, only a measure of accuracy. On Hand-foot Reaction (14) there is the additional speed measure reflecting the time taken to learn the combination between stimuli and controls.

In this table, only ten of the twenty-eight coefficients are significant at better than 5% level. The mean intercorrelation between the four tests, irrespective of whether the measures are speed or accuracy, is 0.12; the mean intercorrelation between all accuracy measures, 0.10; and that between all speed measures, 0.10. There is no suggestion of a general sensory-motor factor or group factors of speed and accuracy. Of the ten significant correlation coefficients, seven are between speed on one test and accuracy on another, two between speed measures, and one between accuracy measures. With the exception of the coefficient of -0.17 between Steadiness (Speed) and Two-hand Co-ordination (Accuracy), all the significant correlations are positive, but, for the most part, small, the three largest being between A.L.C. (Accuracy) and Two-hand Co-ordination (Speed), 0.40; A.L.C. (Accuracy) and Hand-foot Reaction (Reaction Speed), 0.30; Two-hand Co-ordination (Speed) and Hand-foot Reaction (Reaction Speed), 0.26.

The main point illustrated by Tables 1 and 2 is that the mental tests tend to intercorrelate quite substantially, whereas the sensory-motor tests show a comparatively high degree of specificity. However, as Vernon (70, p.96) has observed in connection with work on other sensory-motor tests: "At the same time there is not complete specificity; most of the correlations are positive, though some may be negligible and even negative, especially in high-grade or restricted groups."

Tables 3 and 4 show, respectively, the correlations obtained between speed and accuracy measures on the sensory-motor tests and scores on the mental tests. Most of them are small and positive, the mean coefficient for Table 3 being 0.13, and for Table 4, 0.12. Steadiness (Speed) has almost nothing in common with any of the mental tests but Steadiness (Accuracy) is related to Progressive Matrices, Mental Alertness and Technical and Scientific Information. Hand-foot Reaction

(Reaction Speed), Hand-foot Reaction (Accuracy) and Air-pilots Arm-leg Co-ordinator (Accuracy), have very little in common with the mental tests, and Two-hand Co-ordination (Accuracy) correlates only slightly with Progressive Matrices and the two Technical and Scientific tests. The sensory-motor measures which appear to have the most significant "mental" content are Hand-foot Reaction (Learning Speed) and Two-hand Co-ordination (Speed), the largest correlation being 0.28 between Progressive Matrices and H.F.R. (Learning Speed).

It is curious, that of the two co-ordination tests, namely, Air-pilots arm-leg (A.L.C.) and Two-hand (Moede type), the latter, which would ordinarily be regarded as the less complex, shows more overlap with mental tests than the former. This applies particularly to the speed measure on Two-hand Co-ordination, which relates significantly to all the mental tests except Technical and Scientific Information. It would seem that the comparatively large correlation of 0.40 (Table 2) between Two-hand (Speed) and A.L.C. (Accuracy) is attributable to some common co-ordinative rather than cognitive element.

In his investigation on the relationship between different "motor capacities", Muscio (49) used simple motor tests to minimize the mental element which "would tend to produce some positive intercorrelations, even if none existed between the motor capacities themselves." He was aware, however, that intelligence cannot be entirely eliminated. While the somewhat greater overlap usually found between the more complex sensory-motor tests may be due mainly to some common overlap of g, this is apparently not always the case, as shown by the above correlation between two-hand (Speed) and A.L.C. (Accuracy) and also by an average intercorrelation of 0.25 obtained by Farmer (21) between Choice Reaction, Dotting Machine and Pursuit Meter, none of which correlated to any significant extent with intelligence measures.

The fact that many sensory-motor tests show no significant correlation with mental tests should not be taken to imply that mental factors play no part. A certain minimum of comprehension by the subject is necessary for the performance of all voluntary activities, even the simplest, requiring little more than elementary orientation in space and/or time. In more complex activities there may be the additional need to discriminate quickly and accurately between more than one kind of stimulus and response, master the correct order of a series of operations, and understand causal relations. The Hand-foot Reaction test is of this type, but even here, the correlation with conventional mental tests is by no means large, and one is led to conclude that general mental ability in excess of that necessary for a proper appreciation of what has to be done in the sensory-motor task, affords only a slight extra advantage to the subject.

Mental functions in sensory-motor learning

Granted that man is a psycho-physical organism whose voluntary reactions are extensively dependent on the "dynamic unity of mind and body" stressed by certain psycho-therapeutists like Lamb (39), there is still much to be investigated about the rôle of mental factors in the acquisition of improved sensory-motor skill through learning. Although there is no conclusive proof that general intelligence enters more significantly into practice than it does into initial ability at a particular motor task, improvement is not simply a matter of increased muscular efficiency, there are certain central processes involved which may not be covered by conventional mental tests. Freeman (28) has noted implicit muscular activity during mental work; Buegel (6) found that the introduction of ideational elements in a motor task was beneficial to learning, and the results of Lockhart (40) indicate that in learning the game of bowling supplementary theoretical instruction by motion-picture can contribute significantly.

An interesting approach, which shows considerable promise, is the investigation of the effects of "mental practice". There appears to be evidence that proficiency in certain sensory-motor activities can be increased not only be actual performance but also by intensive thinking about them. Two similar experiments, the one involving dart-throwing and the other, basket-ball shooting, were conducted by Vandell, Davis and Clugston (68) on two different samples. The subjects in each were equated for intelligence, skill and other factors and then divided into three groups. On the first day all groups were given physical practice and the one was then dismissed. During the subsequent 18 days one of the remaining groups carried on with a daily stint of physical practice and the other spent 15 minutes daily "mentally throwing" darts or basket-balls at imaginary targets. On the 20th day all three groups were re-tested. The findings indicated that, while no improvement followed when there was no practice of either kind, mental practice was almost as effective as physical practice in producing improvement.

Twining (66) did a comparable study in which three groups of subjects threw rope rings at a peg. On the first day all groups performed physically, and during the following 20 days one group had no practice of any kind, one had regular physical practice, and one spent 15 minutes daily at a mental rehearsal of the first day's activities. After a re-test of all groups on the 22nd day, it was found that the first had not improved significantly; the second, with physical practice had improved 137%; and the third, with mental practice, 36%. Later studies on the effect of mental practice by Clark (10) and Start (62) have tended to confirm its significance in learning a motor skill. Start (62) noted that the degree of improvement attained through the mental practice was unrelated to scores on intelligence tests, which would be in agreement with the view that in physical skills there can be mental concomitants of a kind not necessarily covered by such tests.

Sometimes the mental processes required in a sensory-motor task are so preponderant that it is quite possible to design a special "pencil-and-paper" substitute that successfully incorporates many of the principal components. This was done by Nesberg and Smith (50) who evolved a "pencil-and-paper" test that extensively reproduced results obtained with the Vector Complex Reactometer, a complex reaction-time apparatus used in air-pilot research.

Nature of the test sample

A point which does not appear to have been generally known to earlier investigators is that the degree of relationship among sensory-motor tests and between these and intellectual tests can be markedly influenced by the composition of the sample examined. In a selected high-grade group of air-pilot candidates the correlations would tend to be considerably smaller than in heterogeneous or low-grade groups where general intelligence or cultural differences account for much more of the variance. On a sample of African mine-workers the mean intercorrelation between tests of assembly, sorting, co-ordination and reaction-time was of the order of 0.35. Vernon (70, p.131) has noted that among college students .... "it is often quite difficult to establish a g."

A large battery of sensory-motor tests, three complex and fourteen simple, administered to students by Perrin (51) yielded intercorrelations that for the most part were insignificant and in many cases, about zero. On the other hand, Attenborough and Farber (3) who tested defective boys, obtained a mean intercorrelation of about 0.40 between simple tests like Tapping, Peg-insertion and Nuts-and-bolts assembly; and a mean correlation of about 0.50 between these tests and the Stanford-Binet and Otis Primary intelligence scales. Malpass (42) found that among retarded children of both sexes there was a high degree of relationship between motor proficiency and intelligence, but hardly any among normal children. An investigation on a group of senile patients by Eysenck (20) revealed a mean r of approximately 0.40 between Steadiness, Aiming and Ergo-graph, and a mean r of 0.26 between a Matrices test of intelligence and Steadiness, Ergograph, Dynamometer and Tapping.

It is possible that sensory-motor abilities are linked with g via a major practical or spatial-mechanical group factor k:m identified by Vernon (69, 70), Cox (11, 12) and others. Vernon (71) considers that the most clear-cut distinction is between this group factor and verbal abilities.

Admittedly correlations tend to be very low, but even here there is evidence of some connection. For example, Uhler (67) investigated two samples of children, one with high psychomotor dominance and the other with high verbal dominance, and found that a wide disparity between verbal and psychomotor "intelligence" was rare and associated with some hereditary mental abnormality in the family history.

Grouping of Sensory-motor Abilities

The conclusion reached by Perrin (51), Muscio (49), Seashore (54), Akroyd (1) and others that there is probably no general factor of sensory-motor ability, is extensively supported by subsequent findings, and much of the later basic research in this field has been directed towards the identification of certain principal dimensions in the form of group factors.

The work of McCloy, E. (46), McCloy, C.H. (44,45), Wendler (72), Carpenter (8), Hempel and Fleishman (37) and others has shown that there is considerable overlap between gross motor abilities of the athletic sort, and clear-cut group factors such as strength, flexibility, gross body co-ordination, velocity, dead-weight and energy-mobilization have been identified. However, hardly any relationship has been found between gross motor activities and fine motor skills.

Reymert (52) was one of the earliest to demonstrate that in simple reaction-time experiments there is a high degree of relationship between various combinations of stimulus and muscular response. Earle and Gaw (19) applied two groups of tests, the first requiring mainly speed, and the second, accuracy of movement. Average intercorrelations were 0.36 in the "speed" group, 0.29 in the accuracy group, butonly 0.13 between the two. Results from a battery of 21 practical tests (19 sensory-motor, one spatial relations, and one mechanical assembly) were subjected to factor analysis by Seashore, Buxton and McCollom (56). One of the six factors that emerged they identified five, as follows: (1) Speed of single reaction; (2) Finger-hand speed in restricted oscillatory movements; (3) Forearm and hand speed in oscillatory movements of moderate extent; (4) Steadiness; (5) Manipulation of spatial relations.

A broad motor co-ordination factor was identified in tests like Pursuit, Aiming, Complex serial reaction, and certain dexterity tests, which were applied in the American-Air forces during the last war (Melton, 47). Rather narrow Manual dexterity factors have been found by Cox (12) and Takala (63). Harrell (34) included fifteen manual tests in a battery applied to cotton-mill operatives, and found a strong Manual dexterity factor pervading tests like Block-packing, Dotting, Pin-board and Nuts-and-bolts assembly. In an analysis of test data obtained on radio-assemblers, Goodman (30) and Chapman (9) located a Manual factor in tests of Tracing, Tapping and Dotting. The findings of Fleishman and Hempel (24, 37) that Manual and Finger dexterity are separate factors, is in conformity with those of the Division of Occupational Analysis, as reported by Vernon (70, p.98).

The results of Seashore and his collaborators (55, 56, 57, 60), Reymert (52) and Campbell (7) show that the boundaries of group factors of fine sensory-motor activities particularly those involving simple single-action and serial-

-action responses, cut across the boundaries of specific musculature and specific sense modalities. Seashore, Buxton and McCollom (56) found a lack of relationship between apparently similar activities like visual reaction-time with the right hand and the maximal rate of tapping with the right hand on the same key. Here, precisely the same musculature was employed, but to produce two distinct patterns of movement. In another investigation, where the movements were all single reactions, Seashore and Seashore (60) found that the subject who reacts quickly with his right hand will also tend to do so with his left hand, either of his feet, or his jaw, elbow, shoulder or neck; and the converse holds for the subject who reacts slowly.

In an experiment by Campbell (7) the subjects had a baton strapped to either hand and foot and made a series of responses to a light by striking a metal plate or key. The very large correlations obtained between the respective performances of the different limbs, led to the erroneous conclusion that there is a general factor of motor ability. Actually, they indicate that there is a strong group factor in activities that have the same pattern.

The particular sense-modality used appears to have a moderate effect on sensory-motor performance of the simple, unitary sort like pressing a key in response to a light or sound. For both, reliabilities are usually about 0.80, and correlations between the two on normal, adult Europeans, are of the order of 0.45. As a rule, reaction to the sound stimulus is slightly faster (Woodworth and Schlosberg 73, p.16).

Seashore (59) inclines to the view that certain groups of activities are more or less related because they involve similar patterns of movement, and the grouping is, therefore, functional rather than anatomical. This is borne out among some of the writer's tests. On the Code-signal reaction apparatus for mine winch-operators there is a correlation of 0.61 between reaction to a light and a sound. Speed and accuracy in the "steadiness" activity of tracing along slots with a stylus held in one hand correlate respectively, 0.53 and 0.40 with speed and accuracy in the Handlebars test (16), where a similar but more complex tracing task requires two-hand co-ordination and more critical postural control. On the Moede-type co-ordination machine, the making of rapid, compensatory movements with two hands in order to guide a track below a stationary pencil-point, correlates 0.40 with the activity of making a generally similar orthogonal pattern of compensatory adjustment with one hand and both feet to keep a drifting light-spot inside a target area on the Air-pilots' Arm-leg Co-ordinator (A.L.C.). In this instance, the fact that the Moede-type test is self-paced and the A.L.C. of the pursuit or automatically-paced type, is apparently not very significant. Performance on the Roundabout test (18) correlates 0.27 with that on the A.L.C., but zero with any measures on the Hand-foot Reaction test; the latter involving a pattern of serial discrimination responses, and A.L.C., like the Roundabout, being a continuous pursuit test, although the musculature of both hands and feet was operative in all three tests.

Sometimes the functional differences are less obvious. For example, in the Floating Effect co-ordination test (13), there is also an orthogonal pattern of compensatory move-

-ment, but the marked element of inertia between control and stimulus makes this a rather different pattern from that in the A.L.C. and Moede-type tests where there is a more rigid mechanical relationship between control and stimulus. The correlation between either of the latter and Floating Effect is of the order of 0.30. However, achievement on Floating Effect and on the Willemse-board test (61) in which a long sloping board is tilted left or right to guide a rolling marble past various baffles into a pocket, correlate to an extent of 0.50; the element of inertia to be allowed for in controlling moving stimuli being common to both. Here again, a pursuit or automatically-paced test (Floating Effect) correlates substantially with a self-paced test (Willemse-Board), but sometimes a difference in pacing appears to be quite important. For example, a battery of tests for African winch-operators included the Moede-type Two-hand and a Two-hand Pursuit test in which the subject turned the handles on a pair of orthogonal lead-screws to keep a pointer in contact with a small target-stud moving automatically along an irregular path. The pattern of movements required was practically identical in both tests, the only significant distinction being in the matter of pacing, and as the test sample was a fairly heterogeneous one with a preponderance of rather low-grade subjects, a very high degree of overlap could have been expected between the two tasks. Actually, performance on the Pursuit test was found to correlate only to an extent of 0.26 and 0.09 respectively, with accuracy and speed measures on the Moede-type machine.

One likely explanation for the greater influence of pacing here, lies in the difference between test samples. The air-pilot tests were administered to European matriculants and the winch-operator tests to comparatively uneducated Africans, many of whom had experienced only very superficial contact with European culture. Working in a steady, unhurried fashion is characteristic of the tribal African's way of life. When necessity drives, he is quite capable of quick and efficient spurts of activity, but sustained performance under the pressure of having to work at a fairly fast rate is opposed to his outlook and disposition. Among the African subjects, therefore, the compulsion of an automatically-paced element in a test probably introduces more temperamental variance than among European subjects. However, it is quite possible that the pursuit element in the Two-hand Pursuit test is generally more stressful than it is in the A.L.C. or Floating Effect.

Table 5 gives the intercorrelations obtained between Educational level, a mental test and six "simple" sensory-motor tests administered to 125 African adults, 63 males and 62 females. The tasks involved in the tests were as follows:

Form Series (Grant, 31) : In this primarily mental test, each item-row on a board comprised part of a recurrent pattern of geometrical figures varying in shape, size and colour, and the subjects were required to add two appropriate plastic figures to continue the pattern in each.

Tweezer Nuts: A small square pan, with a vertical rod fixed at the middle, rested on the table within easy reach, and by means of a pair of tweezers in his preferred hand the subject transferred metal nuts, one at a time and as quickly as possible from the pan on to the rod.

Tweezer Mirror: Similar to the above, except that the pan and rod rested behind a shielding case on a table and the subject had to view his activity in a mirror.

Two-hand Sticks: With the pan and rod resting before the subject on the table, nuts had to be transferred to the rod by means of two wooden sticks, one in either hand.

Steadiness Stick: With the pan and rod resting on the table several feet away, nuts had to be transferred to the rod by means of a single long wooden stick having a metal pin projecting at right-angles near the far end. The other end of the stick was held in the subject's preferred hand.

Tray Co-ordination: By suitable tilting of a metal tray held in both hands, steel balls had to be manipulated, one at a time, up a narrow ramp, past a baffle, and through a small opening into a receptacle.

Ring Throwing: With the pan and rod resting on the table several feet away, subjects endeavoured to throw rubber rings over the rod.

The mean intercorrelation between all the sensory-motor tests is 0.18 which surpasses the 5% level of significance. Those most intimately related are Two-hand Sticks and Steadiness Stick (0.37). Although the one involves the musculature of both hands, and the other that of one only, there is some basic similarity in the pattern of movements required in "stick" manipulation.

Educational level correlates 0.50 with the mental test (Form Series) but is unrelated to performance on the sensory-motor tests except Steadiness Stick, where there is a small negative correlation.

Form Series correlates significantly with only one of the sensory-motor tests, namely, Tweezer Mirror (0.28). In this connection it is of interest that significant correlations between general intelligence scores and performance at a mirror tracing task have been reported by Harmon and Oxendine (33).

Tweezer Mirror has its largest correlation with Two-hand Sticks (0.33) but it is also related significantly to all the other sensory-motor tests, and has the largest mean correlation of 0.24. The motor conflict induced by the mirror image made this the most complex and "difficult" of the sensory-motor tasks.

Ring Throwing, with the smallest mean correlation of 0.12, is the most specific test, which was expected, as the pattern of movement is very different from the fine, sustained manipulations required in the others.

Tray Co-ordinations (a two-handed task) is most closely related to Tweezer Nuts (0.31) and Tweezer Mirror (0.27) which are both one-handed tasks; but it is unrelated to Two-hand Sticks, notwithstanding the apparently strong common element of two-hand co-ordination. Here again, the respective patterns of co-ordination required, are very different.

Even in this comparatively low-grade sample, the sensory-motor abilities measured by the tests appear to be predominantly specific, and there is no indication of a general factor of dexterity or handiness.

The work of Fleishman and Hempel.

Because of the laborious and costly nature of testing for sensory-motor abilities, which usually requires individual administration and special apparatus, experimental groups and the number of tests applied are often rather small, and factorial results statistically insignificant. Notable exceptions are the large scale investigations done by Fleishman (23, 26, 27) and Fleishman and Hempel (24, 25, 37) which merit particular consideration.

In a preceding review of earlier analyses by himself and others, Fleishman (22) described certain factors which appeared to be the most consistent and distinctive, and others which were only tentative. He stressed that the list was based solely on the results of factor-analysis studies, and acknowledged that there are other ways of classifying and describing motor skills. Briefly, his categorization was as follows:-

(a) MOTOR FACTORS: (1) Reaction Time. Speed of response to a stimulus by a prescribed movement. (2) Tapping. Rate of simple oscillatory movements requiring no accuracy. (3) Manual Dexterity. Ability to make rapid and controlled manipulations with the hand or arm. (4) Finger Dexterity. Ability to make rapid and controlled manipulations with the fingers. (5) Steadiness. Precision and steadiness in making arm-hand positioning movements where strength and speed are minimal. (6) Aiming. Speed and precision in making a series of directed movements requiring eye-hand co-ordination. (7) Motor Kinesthesis or Gross Precision. Somewhat precise postural or bodily adjustments to kinesthetic cues when physical equilibrium is disturbed.

Factors not so well-defined were: (8) Psychomotor Co-ordination. Ability to integrate gross or fine movements of moderate scope. This factor was identified whenever more complex tests were used. Possibly it might be split into several sub-factors. (9) Ambidexterity. Ability of right-handed subjects to work fast with the non-preferred hand. It was identified in only one study. (10) Psychomotor Precision. A very vague factor which may be the same as Manual or Finger Dexterity. Tests on which it is loaded actually appear to require fast rather than accurate working.

(b) NON-MOTOR FACTORS: (11) Spatial Relations. Ability to relate different responses to different stimuli, where either stimuli or responses are arranged in spatial order. (12) Perceptual Speed. Ability to make rapid recognitions and comparisons of visual forms. (13) Mechanical Experience. The only motor test with loadings on this factor was Two-hand Co-ordination (presumably the Pursuit type used in the U.S.A.A.F.). Printed tests with loadings on it were Mechanical Information, Mechanical Principles, Biographical Information, Pilot Technical Vocabulary, and Reading Comprehension. (14) Pilot Interest. This factor was distinct from Mechanical Experience, and included a Rudder Control Test, general information, and previous flying experience.

A large number of apparatus and printed psychomotor tests selected or specifically designed to measure most of the above categories was administered by Fleishman (23) to 400 subjects. Results confirmed the existence of the following factors: (1) Reaction Time; (2) Tapping (called "Wrist-finger Speed); (4) Finger Dexterity; (5) Steadiness; (6) Aiming;

(7) Motor Kinesthesis (called "Postural Discrimination");
(8) Psychomotor Co-ordination; (11) Spatial Relations.

An additional factor emerged, which had not been previously hypothesized. It was defined as Rate of Arm Movement and related mainly to those tests requiring rapid arm movements of a gross kind, like Aiming and Placing.

Of particular interest is the confirmation of "Psychomotor Co-ordination" as a distinct factor. It was most consistently present in Rudder Control, Rotary Pursuit, and various Complex Co-ordination Tests, and was also related to Dynamic Balance. Fleishman (23) broadly defines the factor "as representing either co-ordination of the large muscles of the body, in movements of moderate scope, or co-ordination of such movements with the perception of a visual stimulus."

Fleishman and Hempel (24) analysed data from a battery of fifteen widely used printed and apparatus dexterity tests, including the Purdue Pegboard, O'Conner Finger Dexterity, Santa Ana Dexterity, Minnesota Rate of Manipulation, Tapping, Tracing and Marking. They identified five factors accounting for the range of tasks investigated, namely, (1) Finger Dexterity; (2) Manual Dexterity; (3) Wrist-finger Speed; (4) Aiming; and (5) Positioning. The results indicated that the concept of "manual dexterity" as a unitary ability, is untenable.

This study was followed by one in which Hempel and Fleishman (37) investigated a battery of 46 printed, manipulative and gross physical tests. Among the latter, the factors identified were: (a) Strength of limbs and trunk; (b) Flexibility of the legs and trunk. Ability of the muscles to endure and recover from strain and distortion. (c) Balance; (d) Gross body co-ordination; (e) Energy mobilization. Ability to mobilize quickly and effectively a maximum of energy or force. In the "manipulation" group the main factors were: (a) Manual Dexterity; (b) Finger Dexterity; (c) Arm-hand Steadiness; (d) Aiming. Proficiency in the gross physical tasks was found to be quite independent of that in the fine manipulative ones.

Out of a very extensive battery administered to over 1000 naval pilot candidates, Fleishman and Hempel (25) selected and analysed the intercorrelations between sixteen apparatus psychomotor tests and seven printed tests designed as possible substitutes for apparatus tests. Of the nine factors identified, the following four were confined to the apparatus psychomotor tests: (a) Psychomotor Co-ordination I, emphasizing sensitive and highly-controlled adjustments in movements of restricted scope; (b) Psychomotor Co - ordination II, relating more to the gross co-ordination of several limbs in making larger movements; (c) Manual Dexterity, which more adequately covers the factor previously called "Psychomotor Precision"; (d) Rate Control, restricted to those apparatus tests requiring "the ability to make continuous motor adjustments relating to changes in speed of direction of a continuously moving object."

Four factors measurable by both apparatus and certain printed tests were: (a) Integration, "which required the ability to utilize and apply a number of disparate cues and activities quickly into an integrated resultant response". (b) Spatial Relations I, "the ability to interpret spatial characteristics of the stimulus situation"; (c) Spatial Relations II, concerned with "directional discrimination and orientation of movement patterns"; (d) Visualization, emerged as separate from Spatial

Relations. They also found a Perceptual Speed factor, which was restricted to the printed tests.

From the standpoint of equipment design, Brown and Jenkins (5) had classified motor abilities into (a) Static reactions, where a body member must be held for a period in a definite spatial position; (b) Movement reactions, requiring the ability to make co-ordinated, accurate or rhythmic movements at a certain rate or along specific loci; and (c) Positioning reactions, in which the body member must be moved to a prescribed position in space, the terminal accuracy of the response being the most important feature. Each of these classes were subdivided into more restricted ones like continuous movements and intermittent movements, etc.

To investigate two of the main categories above, namely, Positioning movements and Static reactions, Fleishman (26) tested 200 subjects on a wide range of practical tasks. From the results of the factor analysis he concluded that "skill in static reactions is usefully considered a separate class of skills from positioning movements." He found substantial common variance between static reaction tasks, whereas the positioning tasks tended to be highly specific. The three factors extracted were: (1) Arm-hand Steadiness; (2) Moving arm to certain estimated positions not immediately experienced; and (3) Returning arm to a position from which it was just removed. It would seem, from his data, that Arm-hand Steadiness, the most distinct of the three, cuts across the "static-positioning" boundary, because both "positioning" tasks like Track Tracing and Steadiness Precision, and "static" tasks like Steadiness Tremor and Arm Tremor, gave substantial loadings on this factor.

Fleishman (27) considers that those psychomotor skills which fall under the heading movement reactions, are probably the most numerous and important. In these, skill during the movement is of primary interest, whereas in positioning movements the main feature is final precision, and in static reactions, it is the maintenance of a definite limb position. He administered a large and varied battery of "movement" tests to 204 basic trainee airmen. Factor analysis indicated that movement reaction tasks of the sorts investigated fell into several broad classes: (1) Fine Control Sensitivity; (2) Multiple Limb Co-ordination; and (3) Response Orientation.

A factor named "Rate Control" was more tentative. Three other factors that emerged again were Arm-hand Steadiness, Reaction Time, and Speed of Arm Movement; but they did not contribute to performance in the more complex tests.

Guilford's Scheme

A theoretical survey of psychomotor abilities, including both fine skills and gross bodily movements, has been done by Guilford (32), who used the simple, systematic technique of noting "each factor reported in the literature on a separate card, together with its apparent properties and names of some tests that tend to characterize it," and then arranging the cards to form classes of factors. After numerous revisions, he postulated a classification scheme in the form of a matrix. Columns are represented by seven principal types of motor ability, namely, (1) Strength; (2) Impulsion; (3) Speed; (4) Static Precision; (5) Dynamic Precision; (6) Co-ordination; and

(7) Flexibility. Rows refer to the part of the body involved: (a) Gross; (b) Trunk; (c) Limbs; (d) Hand; and (e) Finger.

The concepts of strength, speed and co-ordination are well-known. Impulsion is defined as "the rate at which movements are initiated from stationary positions," and is therefore synonymous with "Reaction-time", whereas Speed refers to the rate of movements after they have started. Static Precision has to do with the accurate maintenance of certain limb or postural positions, and Dynamic Precision with the accuracy of directed movements. Flexibility designates the freedom with which parts of the body can bend, "or the scope of movements dependent upon particular joints."

Factors involving anatomical complexes other than the five above, as, for example, articulation of the speech organs, would require additional rows in the matrix. Guilford also points out that there is uncertainty whether the Tapping factor should be included under "impulsion", as it also involves the repetition of initiated movements, and it had been demonstrated by Seashore et al (57) that Tapping and Reaction-time scores were unrelated, even when the activities were done with same hand.

The specificity of the Tapping factor has been confirmed by Fleishman (23) and others, and it seems pretty obvious that Guilford's concept of "impulsion" will have to be split into (1) Single reactions to discrete stimuli, and (2) Sustained serial activities.

Guilford's system, which is confined to "motor" factors, is certainly the neatest and the most compact which has been proposed so far, but, as he acknowledges himself, it still requires a great deal of experimental substantiation, and may have to be revised.

Unlike Seashore and his collaborators, Guilford appears to regard the particular musculature employed in motor activities as of considerable importance. This has distinct advantages for providing an overall diagram of the entire sensory motor situation, including both gross and fine activities, but there is some risk that it may tend to obscure significant behavioural distinctions among the fine activities, particularly those relating to <u>patterns</u> of movement. Moreover it has been well established that, at least in the activities of the single or discrete reaction type, there is a high degree of intercorrelation between the various body members, and here the particular musculature is therefore not factorially important. Also, the main behavioural difference between gross and fine motor skill is not primarily a matter of the <u>regions</u> of musculature that come into play, but is concerned more with the strength, amplitude, and precision of movements. While fine skills are commonly restricted to the musculature of the fingers, hands and arms, this is not invariably so. For example, the co-ordinated activity in piloting an aircraft, like that in playing football or tennis, makes demands on a large proportion of the total body musculature, but it requires less physical strength, smaller movements and more precise movements, than such "athletic" activities.

<u>Speed and Tempo</u>

In the majority of self-paced sensory-motor tests, whether of the simple, unstructured kind like Tapping where the need for accuracy is minimal, or of the highly structured kind like Handlebars Co-ordination (16) imposing a strict criterion of accuracy, it is usual to instruct the subject to work as quickly as he can. However, another approach, in which the subject is required to work at his "personal tempo" or preferred rate, has received some attention.

Allport and Vernon (2) recorded the preferred rate in performing forty-five varied tasks, and while there was a slight general tendency for the tests to intercorrelate positively, the main group factors that emerged were Verbal Speed (reading, counting and handwriting); Drawing Speed (with the hand on paper, and with the foot on a black-board); and Rhythmic or Motor Speed (Tapping, or contraction of various muscles).

Fifty-nine tests, including natural or congenial rates of card-sorting, cancellation, body-bending and swinging, tapping, drawing and making symmetrical movements with the respective limbs, were administered to 91 university men by Rimoldi (53). The motor tests of personal tempo yielded three factors: Large movements of the trunk and limbs (e.g. Body-bending, limb-swinging, symmetrical movements of limbs); Small movements (mainly represented by tapping with arms, fingers, toes and heels); Drawing (mainly, drawing circles, squares and lines with hands and feet). Tests of reaction-time both simple and complex requiring performance at the subject's maximum speed, were also included in the battery. All of these grouped on another factor. Rimoldi (53) concludes: "If a general tempo factor exists, its influence is very limited due to the existence of definite clusters of speed...."

The natural or preferred speed of children between thirteen and fourteen years in Drawing, Computing, Speaking, Walking and Eating, was investigated by Takala and Partanen (64). They obtained a mean intercorrelation of 0.29 for the boys, and 0.47 for the girls. On Speaking, Walking and Eating only, the mean coefficient for the boys was 0.28, and for the girls, 0.42. The scores were the average values of peers' ratings on a five-point scale. These investigators appear to be in favour of the hypothesis that "in tasks which are strongly overlearned and which occur frequently in daily life" there is a fairly broad "personal tempo" factor which can be predicted to some extent by subjective observation.

The following tests, to be performed in the subjects' own time, were included by Biesheuvel and Pitt (4) in a battery for studying the relationship between personal tempo and the Primary-secondary function variable of temperament: Tapping on a key, Walking, Talking, Writing, Adding, Sorting, Pursuit (Tracing along sinuous printed lines), Drawing crosses on squared paper, Repeated Letters and Identical Pictures. Other tests were: Tapping at maximal rate, Progressive Matrices, and EEG Alpha-frequency. The battery as a whole gave a multiple correlation of 0.61 with the criterion of Primary-secondary function, which was found to be principally related to two kinds of speed tests:
(1) Perceptual: (Repeated Letters, 0.34; Pursuit, 0.39)
(2) Motor: Tapping, fast, 0.31; Crosses, 0.40.

From an inspection of their correlation table, it appears that Walking correlates significantly with only two other variates, namely, Adding (0.29) and EEG Alpha-frequency (-0.33). Rimoldi (53), Harrison (36) and Harrison and Dovens (35) likewise, have found very little connection between walking and other activities. Tapping also, appears to be largely unrelated to most of the other tests, even those which could be expected to have an appreciable motor content. The only exception is a coefficient of 0.39 between Tapping (fast) and Crosses which also involves a rhythmic, serial pattern of movement. However, Tapping correlates more significantly with EEG Alpha-frequency than does any one of the other tests, the coefficients being 0.40 for maximal, and 0.38 for preferred rate. These results afford supporting evidence for the distinctness of the Tapping factor, and also suggest that this kind of rhythmic, serial activity has a strong neurological basis. In view of the correlation of 0.31 between Tapping (fast) and Primary-secondary function, and that of 0.46 between EEG Alpha-frequency and Primary-secondary function, it may also reflect a personal characteristic with quite a wide temperamental significance. A study on preferred Tapping and Metronome rates by Frischeisen-Köhler (29) showed that individual tempo on these tests tends to be very constant in spite of environmental changes, and a comparison of the rate-preferences of twins, siblings and unrelated persons, pointed to the likelihood of some hereditary conditioning.

In line with the results of Allport and Vernon (2) who identified distinct tempo group factors of Tapping and Writing, an average intercorrelation of 0.55 between Writing, Crosses and Pursuit (Line tracing) in the table of Biesheuvel and Pitt (4) suggests a fairly broad "graphological" tempo factor.

As regards the relationship between maximal and preferred rate of work at the same task, Biesheuvel and Pitt found that in the Cross-drawing test there was hardly any difference between the numbers drawn at the preferred and speeded rates. Between Tapping (Preferred) and Tapping (Fast) they obtained a significant correlation of 0.32. Harrision (36) has reported the following correlations between preferred and maximal rates: Card-sorting, 0.17; Tapping, 0.35; Writing, 0.46; Reading aloud, 0.70; Turning a hand-crank, 0.09; and Patting knees with hands, 0.36. (Mean r for all, 0.37); Mangan's (43) table shows correlations of 0.45 and 0.50 between fast working and natural tempo at copying paragraphs. On the writer's Handlebars Co-ordination test (16) there was a correlation of 0.77 between Trial 2, where the need for speed was emphasized in the instructions to the subjects, and Trial 3, where they were told to work at their most natural and congenial rate.

It would appear, that although the extent of overlap between maximal and preferred rate can vary for different tasks, it is often quite substantial, and much the same specific speed factor is operative in both. However, the imposition of maximal rate may sometimes introduce some significant element of temperamental stress. Correlations between maximal and preferred rate are very likely boosted through the fact that although subjects are instructed to work at their natural, preferred pace, they are usually quite aware that their performance is being timed, and they would therefore be stimulated to some extent by an implied urgency. Even when the need for extremely cautious working is emphasized in the instructions and subjects are told positively not to try to work fast, this influence seems to be operative. For example, Trial I on Handlebars, preceded by such instructions, correlates 0.87 with Trial 2 where speed is stressed.

Speed and Accuracy

When a task requires both speed and accuracy, is there a tendency for the fast workers to be less accurate than the slow ones, and vice versa? Thorndike (65) gave a test of arithmetical addition to students, and found that the faster performers were also the more accurate. On the other hand, various mental tests applied by Mudge (48) indicated that there was very little connection between speed and accuracy; a few of the quickest subjects were also the most accurate, and a few of the slowest, the most inaccurate, but among the majority there was no significant relationship. Himmelweit (38) gave four mental tests and a motor test of Track-tracing to a group of neurotic patients, and found no relationship between speed and accuracy on the mental tests, but a significant negative correlation between these measures on the Track-tracer. The results of Longstaff and Porter (41), who investigated the speed and accuracy of pyschology students in answering multiple-choice questions on general psychology, revealed "very little, if any, relationship between the times required and the scores made in answering tests of the type herein used." Mangan (43), who studied speed and other measures in relation to temperament, obtained the following speed/accuracy correlations on three sensory-motor tests; Dotting, -0.07; Measurement, 0.02; and Copying, 0.04.

Speed/accuracy correlations obtained by the writer on six tests are given in Table 6. Of these, one (two-hand co-ordination, Moede-type) is practically zero, and five are

TABLE 1

Intercorrelations between Mental tests

(Decimal points omitted)

N = 161	1	2	3	4	5	6
1 Progressive Matrices (Raven's)						
2 Mental Alertness	56					
3 Arithmetic	34	51				
4 Mathematical Achievement	35	48	57			
5 Tech. and Scientific Information	40	49	16	34		
6 Tech. and Scientific Comprehension	36	49	25	35	57	
7 Aviation Interest Inventory	31	34	09	24	63	47

All coefficients, except 0.09, significant at better than the 5% level.

TABLE 2

Intercorrelations between Sensory-motor tests.

(Decimal points omitted)

	N = 161	1	2	3	4	5	6	7
1	Air-pilots arm-leg Co-ord. (Accuracy)							
2	Steadiness (Speed)	16s						
3	Steadiness (Accuracy)	22s	-10					
4	Hand-foot Reaction (Learn. Speed)	14	-08	19s				
5	Hand-foot Reaction (Reaction Speed)	30s	04	11	10			
6	Hand-foot Reaction (Accuracy)	02	-04	10	06	-14		
7	Two-hand Co-ord. (Accuracy)	14	-17s	07	20s	09	02	
8	Two-hand Co-ord. (Speed)	40s	08	21s	20s	26s	12	03

"s" implies significance at better than the 5% level.

TABLE 3

Correlations between Mental tests and Speed

measures on Sensory-motor tests

(Decimal points omitted)

N = 161	Stead. Speed	H.F.R. Learn. Speed	H.F.R. React. Speed	Two-hand Speed	Mean r
Progressive Matrices	-03	28s	11	24s	15s
Mental Alertness	-02	25s	13	27s	16s
Arithmetic	06	16s	14	16s	13s
Maths. Achievement	04	15	14	23s	14s
Techn. & Scientific Inform.	01	09	-04	14	05
Tech. and Scientific Comp.	-09	17s	17s	22s	12s
Aviation Interest Inventory	00	14	14	18s	12s
Mean r	00	18s	11s	21s	

Mean r of all coefficients = 0.13. Range -0.09 to 0.28

"s" implies significance at better than the 5% level.

TABLE 4

Correlations between Mental tests and Accuracy measures on Sensory-motor tests.

(Decimal points omitted)

N = 161	Air-pilots A.L.C.	Stead. Acc.	H.F.R. Acc.	Two-hand Acc.	Mean r
Progressive Matrices	14	25s	15	19s	18s
Mental Alertness	11	23s	21s	14	17s
Arithmetic	03	04	04	04	04
Maths. Achievement	09	12	08	13	11s
Tech. and Scientific Information	03	21s	11	16s	13s
Tech. and Scientific Comp.	17s	02	15	20s	14s
Aviation Interest Inventory	13	14	04	11	11s
Mean r	10s	14s	11s	14s	

Mean r of all coefficients = 0.12 Range 0.03 to 0.25

"s" implies significance at better than the 5% level.

TABLE 6

Correlations between Speed and Accuracy measures on Sensory-motor tests

Tests	r	N	P
Steadiness (Manual)	-0.10	161	0.204
Handlebars Co-ordination	-0.39	124	0.001
Variable C-ordination	-0.33	47	0.025
Hand-foot Reaction	-0.14	161	0.100
Two-hand Co-ordination (Moede type)	0.03	161	0.704
Dots-cancellation *	-0.26	40	0.103

* Mainly perceptual.

Summary and Conclusions

(1) Sensory-motor activity refers to the physical movements of living organisms, which are never purely mechanical. Psychologists, educationists and employers of personnel are mainly interested in the behavioural aspects of directed movements. The most fundamental directed movements are those of a gross kind like walking, running, jumping etc., in which man is inferior to many other species. However, he is superior to most in activities requiring spatial precision.

(2) Although it is useful to make a broad distinction between gross and fine activities, all have the quantitative facets of strength, speed and amplitude of movement, and the qualitative ones of accuracy, action and pacing. Much of the sensory-motor activity of man is concerned with making things and using them.

(3) It is likely that even in early times certain basic facts about human ability were quite well known, such as the superiority of some individuals over others, the comparative scarcity of the extremely superior and the extremely inferior, the specificity of certain activities, and improvement through practice. In later eras, many homely truths were obscured or ignored by elaborate philosophical speculation, and only achieved "scientific" recognition after the advent of empirical research.

(4) Mental tests tend to intercorrelate quite substantially, whereas sensory-motor tests whow a comparatively high degree of

specificity. Correlations between Mental and Sensory-motor tests are usually very small and positive. The larger relationships sometimes found between more complex sensory-motor tasks may be due to an overlap of _g_ and/or common perceptual and motor patterns.

(5) While the correlations between Sensory-motor tests and conventional Mental tests often appear negligible, results from "mental practice" and other approaches suggest that sensory-motor skills may nevertheless have mental concomitants; albeit of a kind rather specific to the particular task.

(6) Relationships among tests can be influenced by the composition of the test sample. In selected high-grade groups they are usually smaller than in heterogeneous or low-grade groups where intellectual and cultural factors play a greater part.

(7) Fairly large intercorrelations are usual between gross sensory-motor activities, and broad, well-defined group factors have been identified, but there appears to be very little connection between gross and fine skills.

(8) There is probably no general factor of sensory-motor ability; group factors are usually much narrower than those in intellectual fields, and specific factors tend to predominate. The theory, originally proposed by R.H.Seashore, that relationships between sensory-motor activities depend more on the patterns of movement than on the musculature or sense-modalities

employed, is supported by the results of the writer and others.

(9) Among those sensory-motor group factors which have been identified in more than one investigation are:- (a) Reaction-time or Impulsion: Speed in initiating reaction to sensory stimulation. (b) Rapidity of arm-hand movement: Rapid, discrete movements of moderate extent, precision being minimal. (c) Aiming: Rapid, discrete movements of a precise kind requiring eye-hand co-ordination. (d) Tapping: Finger-hand speed in small oscillatory movements. (e) Steadiness: Eye-hand co-ordinations where precision or fine muscular control is emphasized and speed and strength are minimal. (f) Manual dexterity: Rapid and controlled manipulations of a more or less continuous pattern involving mainly the hand. (g) Finger dexterity: Rapid and controlled manipulations of a more or less continuous pattern involving mainly the fingers. (h) Co-ordination: Directed muscular activity requiring the integrated movements of several limbs and/or other body members. (i) Body-Balance: General postural adjustment and the maintenance of equilibrium. (j) Spatial relations: The pairing of responses to stimuli when either set is arranged in a certain spatial order; the judgment of distance in relation to movement.

(10) The extent of overlap between maximal and preferred rate of work can vary for different tasks, but it is often quite considerable. While there appears to be no general factor of "personal tempo", it is likely that group factors of this kind pervade certain clusters of activity, the limits and content of which have still to be defined.

(11) In sensory-motor tasks, correlations between speed of work and accuracy are sometimes zero, but usually slightly negative, and more substantial in complex tasks requiring precise muscular control.

<u>AN APPRAISAL OF SENSORY-MOTOR ACTIVITIES</u>

Abstract

The voluntary muscular movements of man are numerous and diverse. Those of a gross athletic kind where the accent is on the mobilization of physical energy, tend to intercorrelate quite highly, whereas others, requiring smaller and more precise co-ordinations are, for the most part, very specific. To a varying extent, intelligence enters into all sensory-motor skills, although in some the mental accompaniment may be of a rather specific kind not covered by conventional mental tests. The degree of overlap between sensory-motor activities, whether of the structured or unstructured sort, depends more on patterns of movement than on the musculature or sense-modalities employed. Correlations among sensory-motor tests and between these and mental tests are influenced by the composition of the test sample. There appears to be no general sensory-motor factor, but the existence of rather narrow group factors has been established with reasonable certainty. It is likely that there are also broader group factors of "personal tempo."

REFERENCES

1. AKROYD, F. Some tests and correlations of hand-eye co-ordination. Forum Educ., 1928, 6, no.2, 127 - 143.

2. ALLPORT, G.W. and P.E. VERNON. Studies in Expressive Movement. New York, Macmillan, 1933.

3. ATTENBOROUGH, J. and M. FARBER. The relation between intelligence, mechanical ability, and manual dexterity in special school children. Brit J. educ. Psychol., 1934, 4, 140 - 161.

4. BIESHEUVEL, S. and D.R. PITT. The relationship between secondary function and some aspects of speed and tempo of behaviour. Acta psychol., 1955, 11, 373 - 396.

5. BROWN, J.S. and W.O. JENKINS. An analysis of human motor abilities related to the design of equipment and a suggested program of research. In: FITTS, P.M. ed. Psychological Research in Equipment Design. A.A.F. Aviation Psychology Program Research Reports, no. 19. Washington, U.S. Govt. Printing Office, 1947.

6. BUEGEL, H.F. The effect of introducing ideational elements into perceptual-motor learning. J. exp. Psychol., 1940, 27, 111 - 124.

7. CAMPBELL, M. The "personal equation" in pursuit performances. J. appl. Psychol., 1934, 18, 785 - 792.

8. CARPENTER, A. An analysis of the relationship of the factors of velocity, strength, and dead weight to athletic performance. Res. Quart. Amer. Ass. Hlth phys. Educ., 1941, 12, 34 - 40.

9. CHAPMAN, R.L. The MacQuarrie test for mechanical ability. Psychometrika, 1948, 13, 175 - 179.

10. CLARK, L.V. Effect of mental practice on the development of a certain motor skill. Res. Quart. Amer. Ass. Hlth phys. Educ., 1960, 31, 560 - 569.

11. COX, J.W. Mechanical Aptitude. Its Existence, Nature and Measurement. London, Methuen, 1928.

12. COX, J.W. Recent researches into the nature of manual skill. Hum. Factor, Lond., 1934, 8, 356 - 366.

13. DE WET, D.R. Co-ordination and floating effect. J. Nat. Inst. Personnel Res., 1959, 8, 28 - 38.

14. DE WET, D.R. A portable hand-foot reaction test. J. Nat. Inst. Personnel Res., 1960, 8, 106 - 116.

15. DE WET, D.R. An improved steadiness apparatus and its validity for air-pilot selection. J. Nat. Inst. Personnel Res., 1960, 8, 122 - 136.

16. DE WET, D.R. Handlebars: a self-paced test of two-hand co-ordination, and some results on air-pilot candidates. J. Nat. Inst. Personnel Res., 1961, 8, 199 - 208.

17. DE WET, D.R. A variable co-ordination test and its potentiality as a gauge of aptitude for airmanship. Psychol. Afr., 1962, 9, 86 - 99.

18. DE WET, D.R. The Roundabout : A rotary pursuit-test, and its investigation on prospective air-pilots. Psychol. Afr., 1963, 10, 48 - 62.

19. EARLE, F.M. and G. GAW. The measurement of manual dexterities. National Institute of Industrial Psychology Report, 1930, 4, p. 88.

20. EYSENCK, M.D. An exploratory study of mental organization in senility. J. Neurol. Neurosurg. Psychiat., 1945, 8, 15 - 21.

21. FARMER, E. A group factor in sensory-motor tests. Brit. J. Psychol., 1927, 17, 327 - 334.

22. FLEISHMAN, E.A. Testing for psychomotor abilities by means of apparatus tests. Psychol. Bull., 1953, 50, 241 - 262.

23. FLEISHMAN, E.A. Dimensional analysis of psychomotor abilities. J. exp. Psychol., 1954, 48, 437 - 454.

24. FLEISHMAN, E.A. and W.E. HEMPEL. A factor analysis of dexterity tests. Personn. Psychol., 1954, 7, 15 - 32.

25. FLEISHMAN, E.A. and W.E. HEMPEL. Factorial analysis of complex psychomotor performance and related skills. J. appl. Psychol., 1956, 40, 96 - 104.

26. FLEISHMAN, E.A. An analysis of positioning movements and static reactions. J. exp. Psychol., 1958, 55, 13 - 24.

27. FLEISHMAN, E.A. Dimensional analyses of movement reactions. J. exp. Psychol., 1958, 55, 438 - 453.

28. FREEMAN, G.L. The spread of neuro-muscular activity during mental work. J. gen. Psychol., 1931, 5, 479-494.

29. FRISCHEISEN-KÖHLER, I. The personal tempo and its inheritance. Character Person., 1933, 1, 301 - 313.

30. GOODMAN, C.H. The Mac Quarrie test for mechanical ability. II. Factor analysis. J. appl. Psychol., 1947, 31, 150-154.

31. GRANT, G.V. Preliminary findings in the development of the Form Series Test. Psygram, 1965, 7, no. 3, 4 - 16.

32. GUILFORD, J.P. A system of the psychomotor abilities. Amer. J. Psychol., 1958, 71, 164 - 174.

33. HARMON, J.M. and J.B. OXENDINE. Effect of different lengths of practice periods on the learning of a motor skill. Res. Quart. Amer. Ass. Hlth phys. Educ., 1961, 32, 34 - 41.

34. HARRELL, W. A factor analysis of mechanical ability tests. Psychometrika, 1940, 5, 17 - 33.

35. HARRISON, R. and R.M. DOVENS. Is rate of voluntary bodily movements unitary? J. gen. Psychol., 1926, 18, 31 - 39.

36. HARRISON, R. Personal tempo and the interrelationship of voluntary and maximal rates of movement. J. gen. Psychol., 1941, 24, 343 - 379.

37. HEMPEL, W.E. and E.A. FLEISHMAN. A factor analysis of physical proficiency and manupulative skills. J. appl. Psychol., 1955, 39, 12 - 16.

38. HIMMELWEIT, H. Speed and accuracy of work as related to temperament. Brit. J. Psychol., 1946, 36, 132 - 144.

39. LAMB, W. Individual movement therapy. Ment. Hlth, Lond., 1953, 13, 16 - 19.

40. LOCKHART, Aileene. The value of the motion picture as an instructional device in learning a motor skill. Res. Quart. Amer. Ass. Hlth Phys. Educ., 1944, 15, 181 - 187.

41. LONGSTAFF, H.P. and J.P. PORTER. Speed and accuracy as factors in objective tests in general psychology. J. appl. Psychol., 1928, 12, 636 - 642.

42. MALPASS, L.F. Motor proficiency in institutionalized and non-institutionalized retarded children and normal children. Amer. J. ment. Defic., 1960, 64, 1012 - 1015.

43. MANGAN, G.L. A factorial study of speed, power and related temperament variables. Brit. J. educ. Psychol., 1959, 29, ~~26~~, 144 - 154.

44. McCLOY, C.H. A preliminary study of factors in motor educability. Res. Quart. Amer. Ass. Hlth phys. Educ., 1940, 11, 28 - 39.

45. Mc Cloy, C.H. The measurement of speed in motor performance. Psychometrika, 1940, 5, 173 - 182.

46. Mc Cloy, E. Factor analysis methods in the measurement of physical abilities. Res. Quart. Amer. Ass. Hlth phys. Educ., 1935, 6, 81 - 85.

47. MELTON, A.W. ed. Apparatus Tests. Army Air Forces Aviation Psychology Program Research Reports, no. 4. Washington, U.S. Govt. Printing Office, 1947.

48. MUDGE, E.L. Time and accuracy as related to mental tests. J. educ. Psychol., 1921, 12, 159 - 161.

49. MUSCIO, B. Motor capacity with special reference to vocational guidance. Brit. J. Psychol., 1922, 13, 57-84.

50. Nesberg, L.S. and K.U. SMITH. Measurement of a complex psychomotor performance by means of a printed test. J. appl. Psychol., 1950, 34, 309 - 312.

51. PERRIN, F.A.C. An experimental study of motor ability. J. exp. Psychol., 1921, 4, 24 - 56.

52. REYMERT, M.L. The personal equation in motor capacities. Scand. sci. Rev., 1923, 2, 177 - 194.

53. RIMOLDI, H.J.A. Personal tempo. J. abnorm. soc. Psychol., 1951, 46, 283 - 303.

54. SEASHORE, R.H. The Stanford motor skills unit. Psychol. Monogr., 1928, 39, no. 2, 51 - 66 (Whole no. 178).

55. SEASHORE, R.H. An experimental and theoretical analysis of fine motor skills. Amer. J. Psychol., 1940, 53, 86 - 98.

56. SEASHORE, R.H., C.E. BUXTON and I.N. Mc COLLOM. Multiple factorial analysis of fine motor skills. Amer. J. Psychol. 1940, 53, 251 - 259.

57. SEASHORE, R.H., R. STARMANN, W.E. KENDALL and J.S. HELMICK. Group factors in simple and discrimination reaction time. J. exp. Psychol., 1941, 29, 346 - 349.

58. SEASHORE, R.H., F.J. DUDEK and W. HOLTZMAN. A factorial analysis of arm-hand precision tests. J. appl. Psychol., 1949, 33, 579 - 584.

59. SEASHORE, R.H. Work and motor performance, pp. 1341 - 1362. In: STEVENS, S.S., ed., Handbook of Experimental Psychology, New York, John Wiley, 1958.

60. SEASHORE, S.H. and R.H. SEASHORE. Individual differences in simple auditory reaction times of hands, feet and jaws. J. exp. Psychol., 1941, 29, 342 - 345.

61. SKAWRAN, R.F. The Willemse board as a personality test. _Psychol. Afr._, 1962, 9, 72 - 85.

62. START, K.B. Relationship between intelligence and the effect of mental practice on the performance of a motor skill. _Res. Quart. Amer. Ass. Hlth phys. Educ._, 1960, 31, 644 - 649.

63. TAKALA, Annika. On the factors of manual skill. _Acta psychol. fenn._, 1951, 1, 108 - 128.

64. TAKALA, Martti _and_ N. PARTANEN. Psychomotor expression and personality study III. The problem of "personal tempo". _Scand. J. Psychol._, 1964, 5, 161 - 170.

65. THORNDIKE, E.L. The relation between speed and accuracy in addition. _J. educ. Psychol._, 1915, 5, 537 - 541.

66. TWINING, W.E. Mental practice and physical practice in learning a motor skill. _Res. Quart. Amer. Ass. Hlth phys. Educ._, 1949, 20, 432 - 435.

67. UHLER, C. Wide span psychometric patterns. _Delaware St. Med. J._, 1937, 9, 1 - 6.

68. VANDELL, R.A., R.A. DAVIS _and_ H.A. CLUGSTON. The function of mental practice in the acquisition of motor skills. _J. gen. Psychol._, 1943, 29, 243 - 250.

69. VERNON, P.E. The structure of practical abilities. _Occup. Psychol._, 1949, 13, 81 - 96.

70. VERNON, P.E. The Structure of Human Abilities. London, Methuen, 1950. 2nd Ed. 1961.

71. VERNON, P.E. The classification of abilities. Educ. Res., 1960, 2, 184 - 193.

72. WENDLER, A.J. A critical analysis of test elements used in physical education. Res. Quart. Amer. Ass. Hlth phys. Educ., 1938, 9, 64 - 76.

73. WOODWORTH, R.S. and H. SCHLOSBERG. Experimental Psychology. New York, Henry Holt, 1954.

SECTION IV

VISUAL PERCEPTION APPARATUS

SECTION IV

Introduction

In all directed sensory-motor tasks, perceptual factors (usually visual, but sometimes also auditory, kinaesthetic, or tactile) play a significant part. Perception is the vital terminal coupling in the sensory chain that links the organism with external reality. Without perception the performance of successful directed activities is impossible, because there can be neither proper awareness of the requirements of the practical situation nor adequate "feed-back" after motor responses to indicate whether they have been successful or not.

In some activities, however, motor responses are of comparatively slight or even negligible importance, and the principal factors are perceptual, as, for example, in radar-screen observation for air-traffic control and interception from ground stations. The first instrument described in this section, a Locus-estimation Test, was specifically designed to test candidates for this vocation, and can be administered to individuals or to a small group of up to six at once.

The second apparatus, a Cancellation Timer, is not a "test" in itself, but actually a basic instrument for the standard administration of various printed Cancellation tests, when it is desirable to obtain a fairly accurate time measure for each line completed. Cancellation provides a simple perceptual-motor activity that has been found useful for measuring "sustained attention" of the kind required in many routine tasks, particularly of a clerical nature. The perceptual element in recognising items that have to be cancelled is more important than the repetitive motor response of making pencil-strokes through them.

Technical details of the Cancellation Timer are followed by an account of a Dots-cancellation test administered to air-pilot candidates by means of this apparatus.

The third apparatus in this section, a Dual-purpose "Fall" Tachistoscope, is likewise, not a test in itself, but a basic device for administering various tests requiring the presentation of either moving or stationary visual stimuli for short periods.

After a full description of the technical aspects, an account is given of tests of Speed of Perception and Span of Attention administered with the apparatus to air-pilot candidates.

Reprinted from the Journal of the National Institute for Personnel Research.

J. NAT. INST. PERSONNEL RES., 1959, 8, 48—55

A LOCUS-ESTIMATION TEST

D. R. DE WET

(Received 24th November, 1958)

The steady advance in Radar technology, particularly that branch concerning the remote observation and control of flying objects from ground stations, has secured it a position of world-wide importance, both as a versatile navigational instrument in peace time and a formidable defensive and aggressive weapon in war.

At civilian aerodromes, during bad weather conditions when physical visibility is insufficient for the pilot of an incoming aircraft to determine his exact bearings, he has to rely on information from the control-tower, where his course and position in relation to the terrain are represented by a moving light-spot or "blimp" on a calibrated screen.

At military stations during periods of national emergency, the radar light-spot serves as a means of identifying the course, position and probable destination of an enemy aircraft or missile, and (time permitting) enables appropriate action to be implemented. This gives rise to a double situation, portrayed on the radar-screen by two moving light-spots, one representing the enemy object or "target," and the other the pursuing "interceptor."

Basing their calculations on this "screen-picture" and making the necessary corrections and modifications as it changes, the ground-control personnel strive to guide the pilot of the interceptor to a coign of vantage from which he can destroy the target most promptly and efficiently.

The precise form of this optimum strategic appropinquity will depend on how the pursuit develops, but in all cases it should be attained as soon and directly as possible, because the interceptor-machine may have little or no superiority over the target in point of speed and the latter could reach its objective before preventive measures are really practicable.

Although lacking much of the traditional "glamour" popularly associated with fighter-pilots, the role of the ground interception staff has come to be so vital in national defence that certain countries have encouraged some of their most experienced pilots and navigators to withdraw from active flying and devote themselves to this section of the Service.

It is also considered well worth while that selection for suitability in this occupation be carried out to the best of available testing resources and knowledge of the job requirements.

Actual radar-screen observation constitutes only one aspect of the activities involved; there are many others, calling for ability in such matters as navigation, administration, electronics, meteorology, aerodynamics, the potentialities and limitations of different aircraft and missile types, and the wide cognate field of tactics in modern aerial combat.

Test Purpose.

It is not the intention here to attempt a detailed enumeration of aptitude tests which have been applied with varying success to cover these diverse facets of ground-control work, or a discourse on the cognitive acuity and dispositional crasis most adapted to this occupation as a whole.

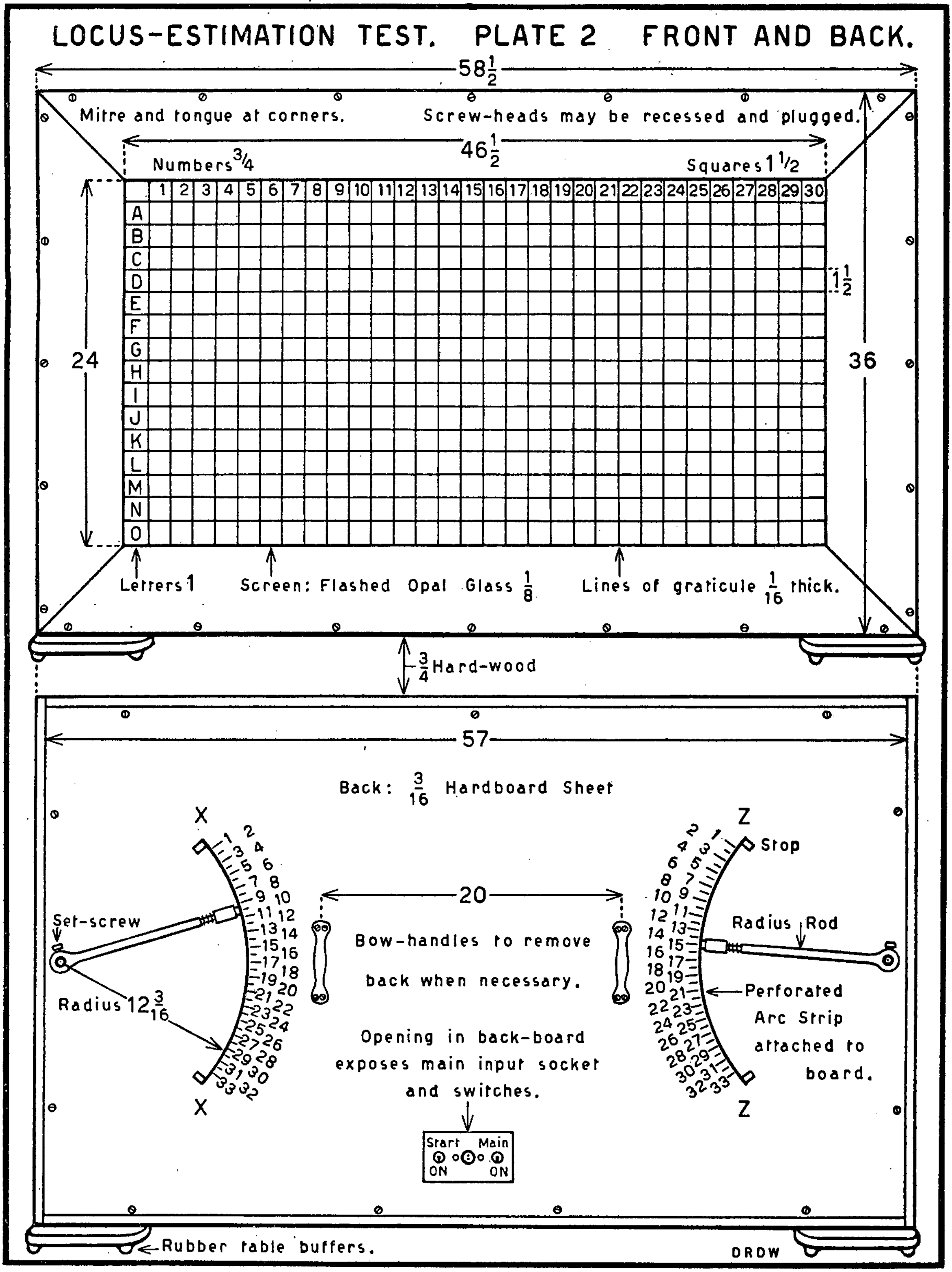
The description that follows relates to a piece of fairly simple equipment designed to incorporate some of the elements more specific in radar-screen observation.

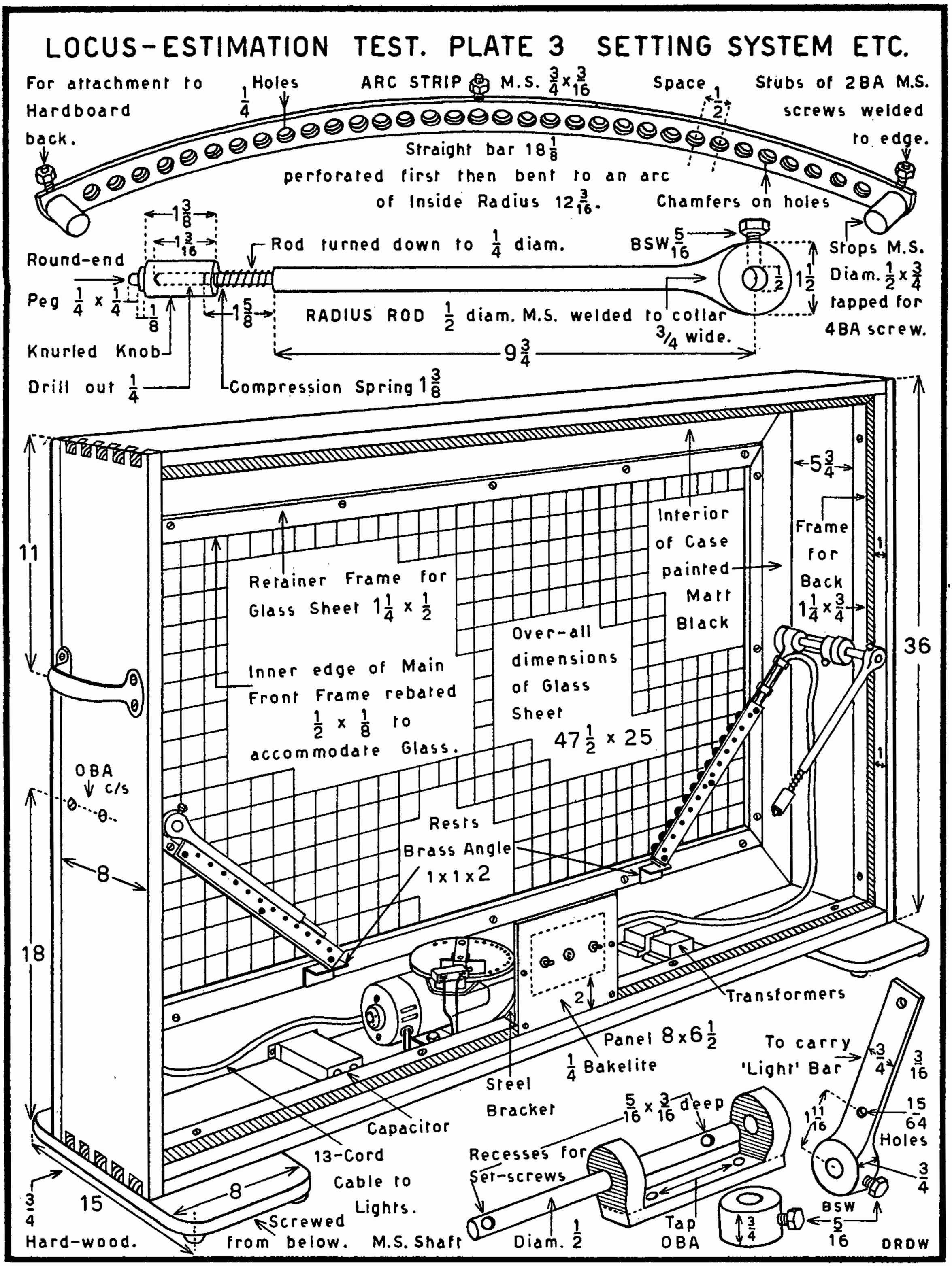
Orientation tests of a static nature have shown validity in this respect; it was considered, however, that the battery might be usefully supplemented by a dynamic test in some ways more closely resembling practical conditions. Three main features included are: Distribution of Attention, Judgment of movement and direction, and the Spatial fixation of bearing-points.

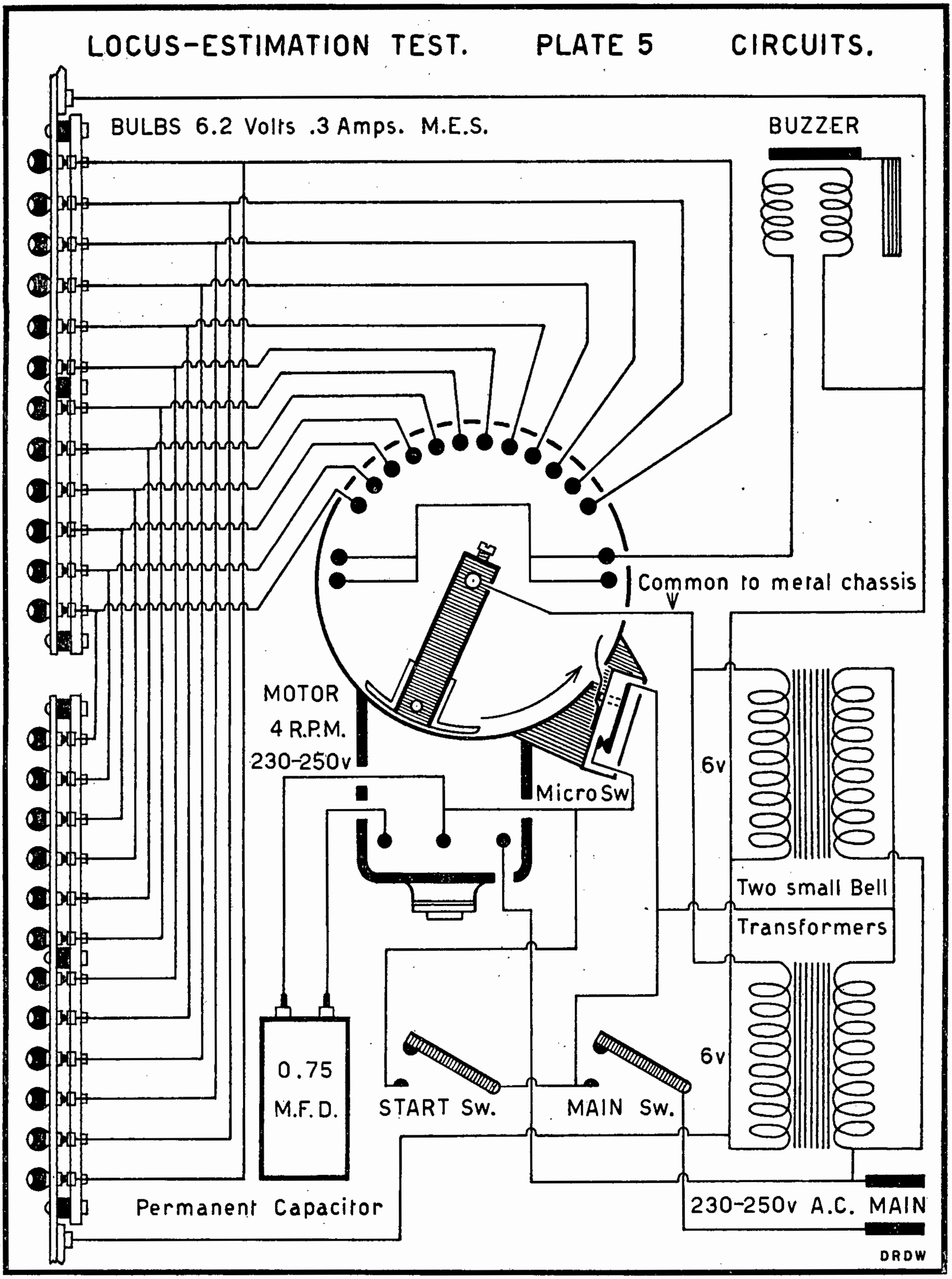
This Locus-estimation Test may be administered to individual subjects or to a small group of up to six at once.

Procedure.

Subjects are seated at a distance of approximately eight feet from a translucent


LOCUS-ESTIMATION TEST. PLATE 2 FRONT AND BACK.
58½
Mitre and tongue at corners.
Screw-heads may be recessed and plugged.
46½
Numbers ¾
Squares 1½
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
A B C D E F G H I J K L M N O
1½
24
36
Letters 1
Screen: Flashed Opal Glass ⅛
Lines of graticule 1/16 thick.
¾ Hard-wood
57
Back: 3/16 Hardboard Sheet
X
Z
Stop
Set-screw
Radius 12 3/16
20
Bow-handles to remove back when necessary.
Radius Rod
Perforated Arc Strip attached to board.
Opening in back-board exposes main input socket and switches.
Start ON
Main ON
Rubber table buffers.
DRDW


LOCUS-ESTIMATION TEST. PLATE 3 SETTING SYSTEM ETC.
For attachment to Hardboard back.
1/4 Holes
ARC STRIP M.S. 3/4 x 3/16
Space 1/2
Stubs of 2BA M.S. screws welded to edge.
Straight bar 18 1/8 perforated first then bent to an arc of Inside Radius 12 3/16.
Chamfers on holes
1 3/8
1 3/16
Round-end Peg 1/4 x 1/4
1/8
Knurled Knob
Drill out 1/4
1 5/8
Compression Spring 1 3/8
Rod turned down to 1/4 diam.
RADIUS ROD 1/2 diam. M.S. welded to collar 3/4 wide.
9 3/4
BSW 5/16
1/2
1 1/2
Stops M.S. Diam. 1/2 x 3/4 tapped for 4BA screw.
11
Retainer Frame for Glass Sheet 1 1/4 x 1/2
Inner edge of Main Front Frame rebated 1/2 x 1/8 to accommodate Glass.
Over-all dimensions of Glass Sheet 47 1/2 x 25
Interior of Case painted Matt Black
5 3/4
Frame for Back 1 1/4 x 3/4
1
1
36
OBA c/s
8
18
Rests Brass Angle 1 x 1 x 2
Transformers
2
Panel 8 x 6 1/2
1/4 Bakelite
Steel Bracket
Capacitor
13-Cord Cable to Lights.
3/4
15
8
Hard-wood.
Screwed from below.
M.S. Shaft
Recesses for Set-screws
Diam. 1/2
5/16 x 3/16 deep
Tap OBA
3/4
To carry 'Light' Bar
3/4
3/16
1 11/16
15/64 Holes
3/4
BSW 5/16
DRDW


LOCUS-ESTIMATION TEST. PLATE 5 CIRCUITS.
BULBS 6.2 Volts .3 Amps. M.E.S.
BUZZER
Common to metal chassis
MOTOR
4 R.P.M.
230-250v
Micro Sw
6v
Two small Bell
Transformers
6v
0.75
M.F.D.
START Sw.
MAIN Sw.
Permanent Capacitor
230-250v A.C. MAIN
DRDW

Locus-estimation compared with other tests applied to ground-control interceptor operators

In addition to the Locus-estimation test, the experimental battery applied to Air-force G.C.I. operators included Mental Alertness, Progressive matrices, Personality Assessment and three pictorial pencil-and-paper tests specifically devised by the Royal Air Force for selection in this vocation. The latter tests, called respectively, "Tables and Graphs", "Angles, Bearings and Directions", and "Ships", presented diagrammatic problems relating to various aspects of elementary navigation.

The intercorrelations obtained between the tests (excluding Personality Assessment) are given in Table 1. Among the five intellectual tests, the correlations are mostly high, and all significant at better than the five percent level. Mental Alertness, (Mean r, 0.57) shows the widest overlap with other tests. Locus-estimation, which is mainly perceptual and the only apparatus test in the battery, is also the most specific, with the smallest Mean r of 0.41. Its closest relationship is with the "Angles, Bearings and Directions" test (r, 0.53) which, to a considerable extent also involved the estimation of loci, but in static situations.

Table 2 gives the correlations obtained between test scores and achievement on the job in terms of the percentage scored by candidates in practical examinations after training had proceeded for two months.

The best individual predictors of occupational success would appear to be the two intelligence tests and the R.A.F. "Tables and Graphs" test, which are all significant at the one percent level. Then follow the R.A.F. "Angles, Bearings and Directions" test and Locus-estimation, respectively significant at the five and seven percent levels. Personality Assessment and the R.A.F. "Ships" test have the smallest coefficients, which do not reach the 10% level of significance.

A multiple correlation of 0.74, significant at the five percent level, was obtained between the tests and the occupational criterion.

In this battery, the two intelligence tests, namely, Progressive Matrices and Mental Alertness, were found to carry major weights (0.416 and 0.375 respectively), probably because the vocation of the ground-control interceptor operator, while "practical" in certain respects, involves many theoretical subjects like General Knowledge, Electronics, Meteorology and Navigation, and is therefore primarily intellectual. The R.A.F. "Tables and Graphs" test, and Locus-estimation had weights of 0.181 and 0.115 respectively. The weights of the other two R.A.F. tests were negative, namely :
"Ships", -0.261; "Angles, Bearings and Directions", -0.005.

Detailed accounts of research done on the G.C.I. operators are given in the following :

(1) KAMFER, L. Selection of Ground Control Interceptors (G.C.I.). A validation study of the N.I.P.R. test battery with Part II of the G.C.I. training course. Internal Report, 1956.

(2) TROMP, J. Selection of South African Air Force fighter controllers. Psychol. Afr., 1962, 9, 67-71.

Reprinted from the Journal of the National Institute for Personnel Research.

J. NAT. INST. PERSONNEL RES., 1959, 8, 48—55

A LOCUS-ESTIMATION TEST

D. R. DE WET

(Received 24th November, 1958)

The steady advance in Radar technology, particularly that branch concerning the remote observation and control of flying objects from ground stations, has secured it a position of world-wide importance, both as a versatile navigational instrument in peace time and a formidable defensive and aggressive weapon in war.

At civilian aerodromes, during bad weather conditions when physical visibility is insufficient for the pilot of an incoming aircraft to determine his exact bearings, he has to rely on information from the control-tower, where his course and position in relation to the terrain are represented by a moving light-spot or "blimp" on a calibrated screen.

At military stations during periods of national emergency, the radar light-spot serves as a means of identifying the course, position and probable destination of an enemy aircraft or missile, and (time permitting) enables appropriate action to be implemented. This gives rise to a double situation, portrayed on the radar-screen by two moving light-spots, one representing the enemy object or "target," and the other the pursuing "interceptor."

Basing their calculations on this "screen-picture" and making the necessary corrections and modifications as it changes, the ground-control personnel strive to guide the pilot of the interceptor to a coign of vantage from which he can destroy the target most promptly and efficiently.

The precise form of this optimum strategic appropinquity will depend on how the pursuit develops, but in all cases it should be attained as soon and directly as possible, because the interceptor-machine may have little or no superiority over the target in point of speed and the latter could reach its objective before preventive measures are really practicable.

Although lacking much of the traditional "glamour" popularly associated with fighter-pilots, the role of the ground interception staff has come to be so vital in national defence that certain countries have encouraged some of their most experienced pilots and navigators to withdraw from active flying and devote themselves to this section of the Service.

It is also considered well worth while that selection for suitability in this occupation be carried out to the best of available testing resources and knowledge of the job requirements.

Actual radar-screen observation constitutes only one aspect of the activities involved; there are many others, calling for ability in such matters as navigation, administration, electronics, meteorology, aerodynamics, the potentialities and limitations of different aircraft and missile types, and the wide cognate field of tactics in modern aerial combat.

Test Purpose.

It is not the intention here to attempt a detailed enumeration of aptitude tests which have been applied with varying success to cover these diverse facets of ground-control work, or a discourse on the cognitive acuity and dispositional crasis most adapted to this occupation as a whole.

The description that follows relates to a piece of fairly simple equipment designed to incorporate some of the elements more specific in radar-screen observation.

Orientation tests of a static nature have shown validity in this respect; it was considered, however, that the battery might be usefully supplemented by a dynamic test in some ways more closely resembling practical conditions. Three main features included are: Distribution of Attention, Judgment of movement and direction, and the Spatial fixation of bearing-points.

This Locus-estimation Test may be administered to individual subjects or to a small group of up to six at once.

Procedure.

Subjects are seated at a distance of approximately eight feet from a translucent

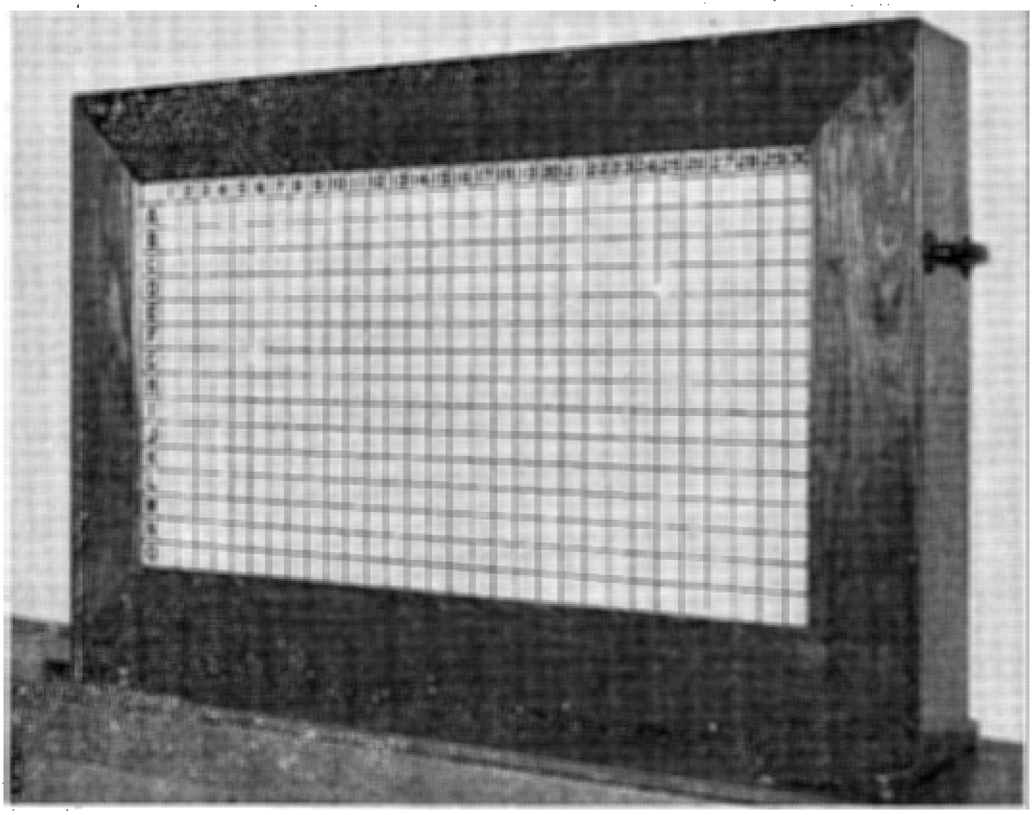


Plate 1 Locus-estimation Test

screen divided up into squares with numerical references along the top horizontal axis and letters down the left vertical axis. (Plates 1 and 2.)

They are each supplied with a board on which to write and an answer sheet having columns of blank spaces labelled one to fifty. The following instructions are then read out:

"This is a test of your ability to judge the spatial and directional relationship between two objects moving at a constant speed but at various angles to each other.

"The objects are represented by two spots of light which will move concurrently across the screen at an angle to each other, one from the right and the other from the left.

"They will travel for a certain distance and then disappear. Your task is to estimate in which square on the screen their respective courses will intersect. You must imagine the lights as continuing in their courses until the one cuts across the course of the other.

The course of each light is always a straight line and the point of intersection is always in some square on the screen.

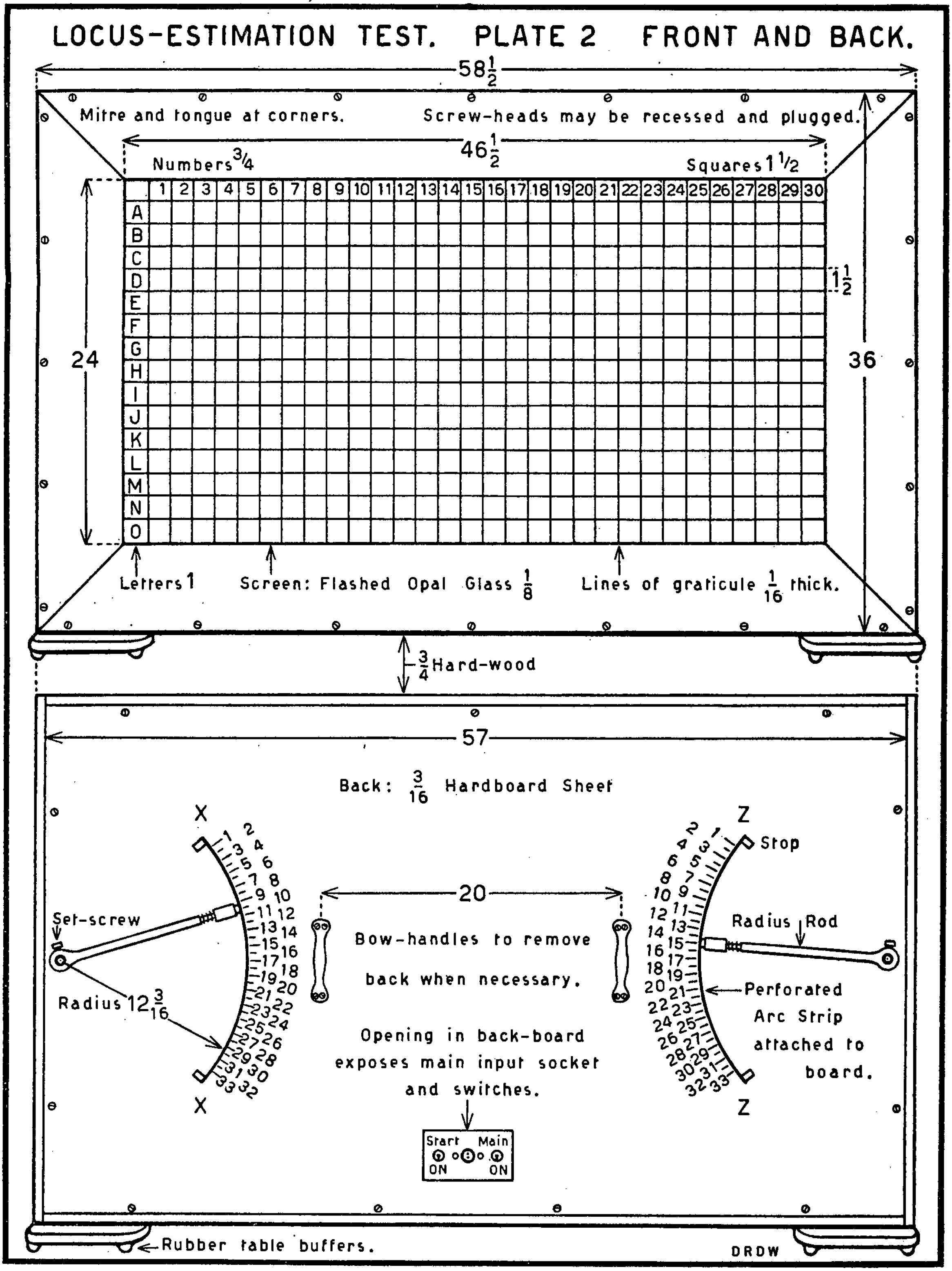
"Squares are identified on your answer sheet by the appropriate letter and number—B4, L13, etc.

A buzzer will sound a warning just before the lights come on, and again after the end of each trial.

You have three trials to illustrate the procedure. Here is the first. (Demonstration and answers given.) N21, B10, J21.

Now we start the test proper. On your answer sheet, as we do each trial, write down the name of the square where you judge the light courses will intersect. The correct square will earn the highest score in each case, but some credit will also be given for answers which deviate within limits."

When the observer is satisfied that all subjects understand exactly what is required, the two radius-rods at the back of the apparatus-case (Plate 2) which were set three times at certain pre-selected positions on their respective arc-strips X and Z for the


LOCUS-ESTIMATION TEST. PLATE 2 FRONT AND BACK.
58 1/2
Mitre and tongue at corners.
Screw-heads may be recessed and plugged.
46 1/2
Numbers 3/4
Squares 1 1/2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
A B C D E F G H I J K L M N O
1 1/2
24
36
Letters 1
Screen: Flashed Opal Glass 1/8
Lines of graticule 1/16 thick.
3/4 Hard-wood
57
Back: 3/16 Hardboard Sheet
X
Z
Stop
20
Set-screw
Bow-handles to remove
back when necessary.
Radius Rod
Radius 12 3/16
Perforated
Arc Strip
attached to
board.
Opening in back-board
exposes main input socket
and switches.
Start
Main
ON
ON
Rubber table buffers.
DRDW

three demonstrations, are again set for the first stimulus presentation of the test series and the apparatus started just as it was for the demonstrations.

While subjects are writing down their answers, the radius-rods are re-set to give the next combination on the observer's list, and when everyone is ready this is presented in the same way; as are the remainder of the fifty items.

Scoring.

Originally credit was given only for identification of the correct square in which the imaginary continuation of the two light-loci would intersect.

It was found, however, that a much better distribution of performance resulted from a system of graded scoring.

Accordingly, three points are given for naming the correct square or "bull's-eye"; two points for any one of the surrounding eight squares; and one point for any one of the outer sixteen squares.

Experimental Finding.

The test was administered to 21 candidates for ground-control interception duties in the S.A.F.

Taking as the criterion a final percentage rating of the pupils by their instructors based on competence in several practical interceptions during training, a Product Moment Correlation of .4 significant at the 7% level, was obtained.

The degree of difficulty experienced in identifying the intersection-points of the various pairs of light-loci, was found to depend considerably on the angle of approach in a given combination. Being easiest when the approach of one locus towards the other most nearly resembled a right-angle, it became most difficult in the nearly "head-on" approach.

A similar pattern of relative difficulty is recognised and accepted by screen observers in actual interception work.

Functioning Principle.

At first various mechanical methods of imparting motion to the light-spots were tried, but with little success. The least unpromising of these was the propulsion of two lamps along rails by long revolving rods with spiral groves; the angles of the two sets of rails and rods being adjustable in relation to each other to give the different approaches.

However, difficulties encountered in maintaining synchronisation of the two stimuli and also in getting them back to zero after each presentation without much delay and fumbling, and the risk of subjects perceiving settings beforehand, led to the abandonment of this method likewise.

Finally a simple *appearance* of motion like that extensively used in the illuminated advertising industry, was found to give satisfactory results. This is the well-known visual Phi-phenomenon, where an illusion of directional movement is imparted to stationary stimuli when they are presented successively in neighbouring positions.

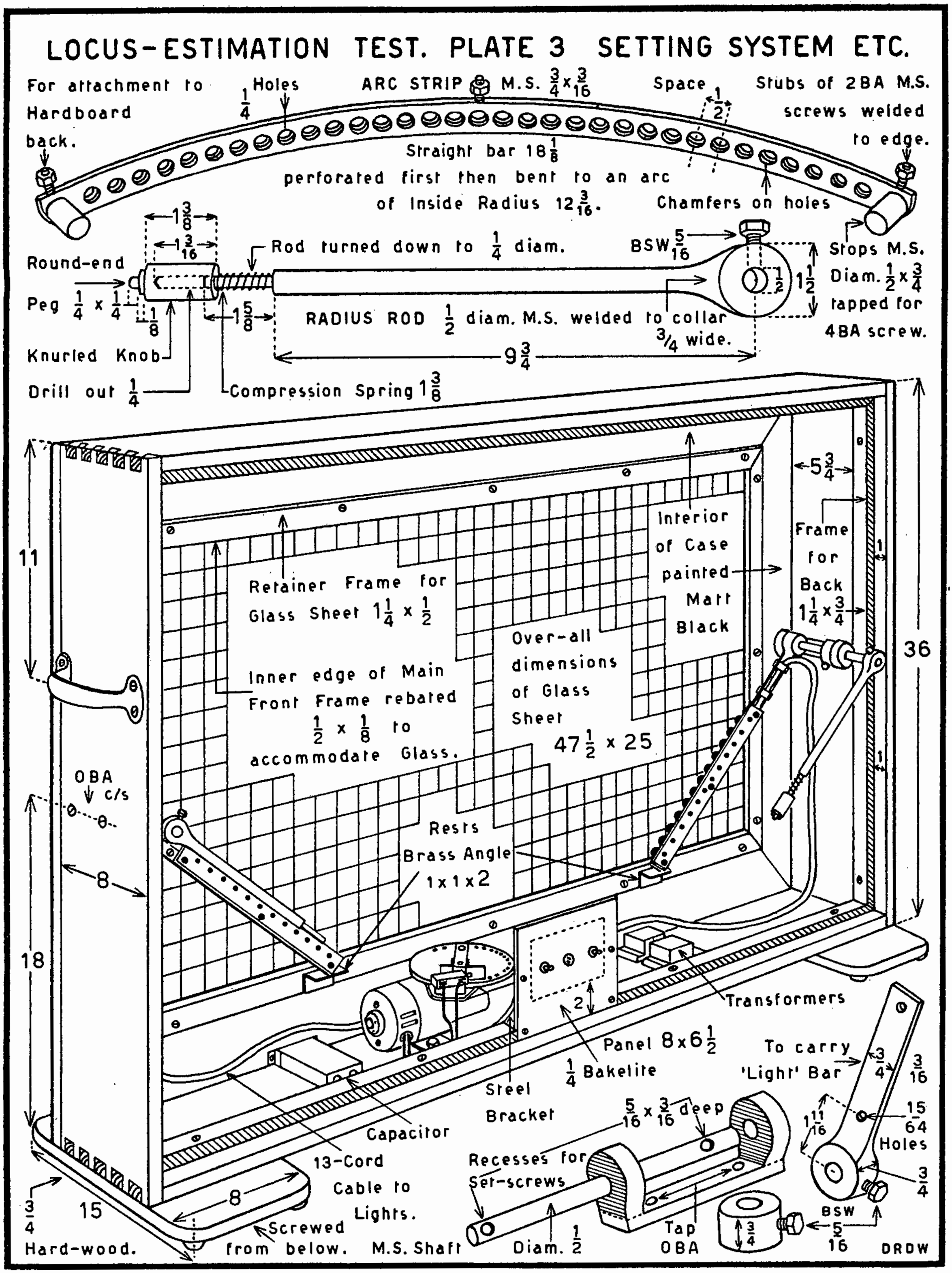
Plates 4 and 5 explain the utilization of this principle in the Locus-estimation Test. Instead of two moving light-spots, there are two bars, each having twelve stationary bulbs mounted upon it at equal distances. Successive illumination of the bulbs by a simple motor-driven distributor produces the semblance of a single light moving along each bar.

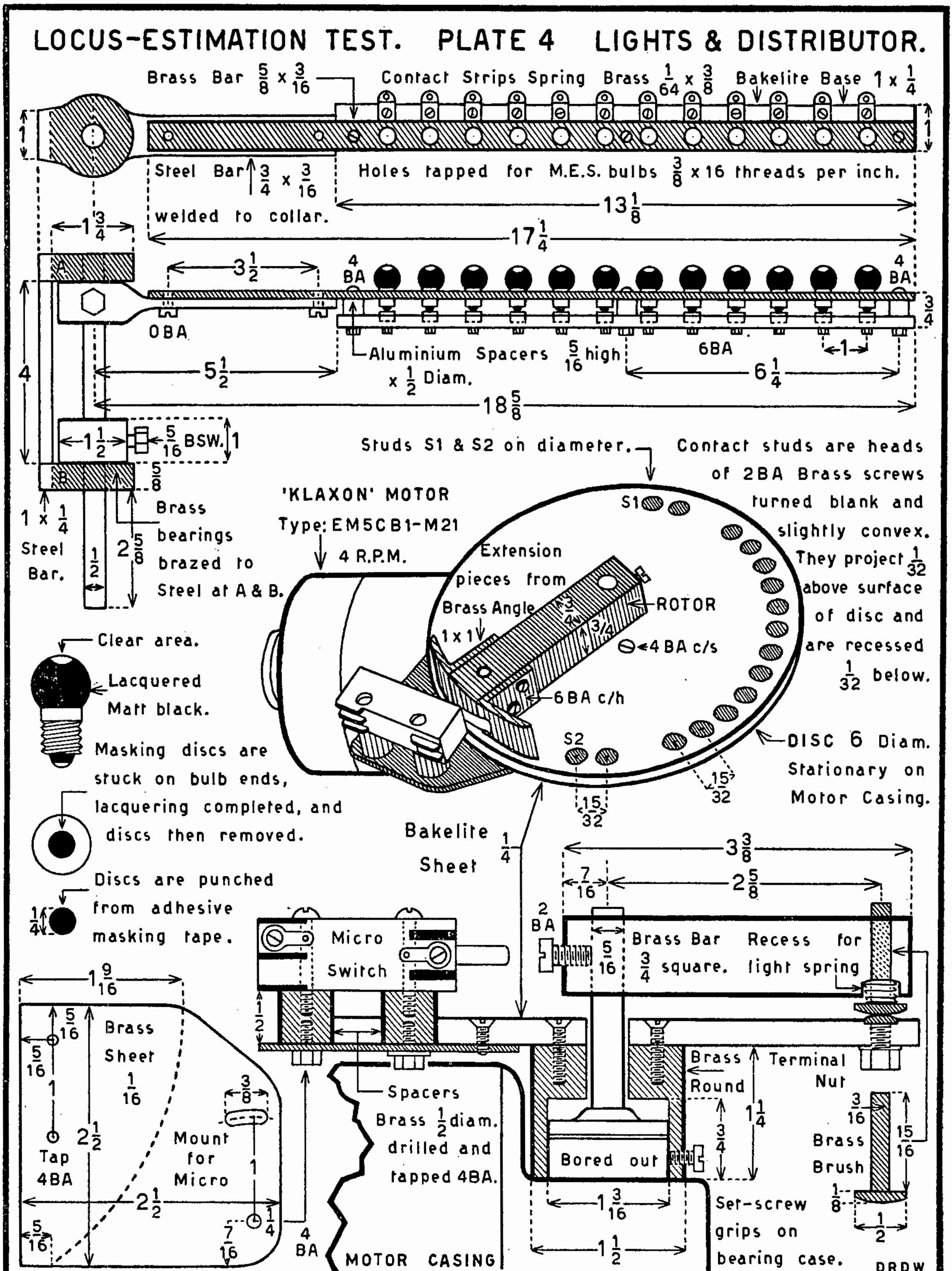
Mechanical Construction.

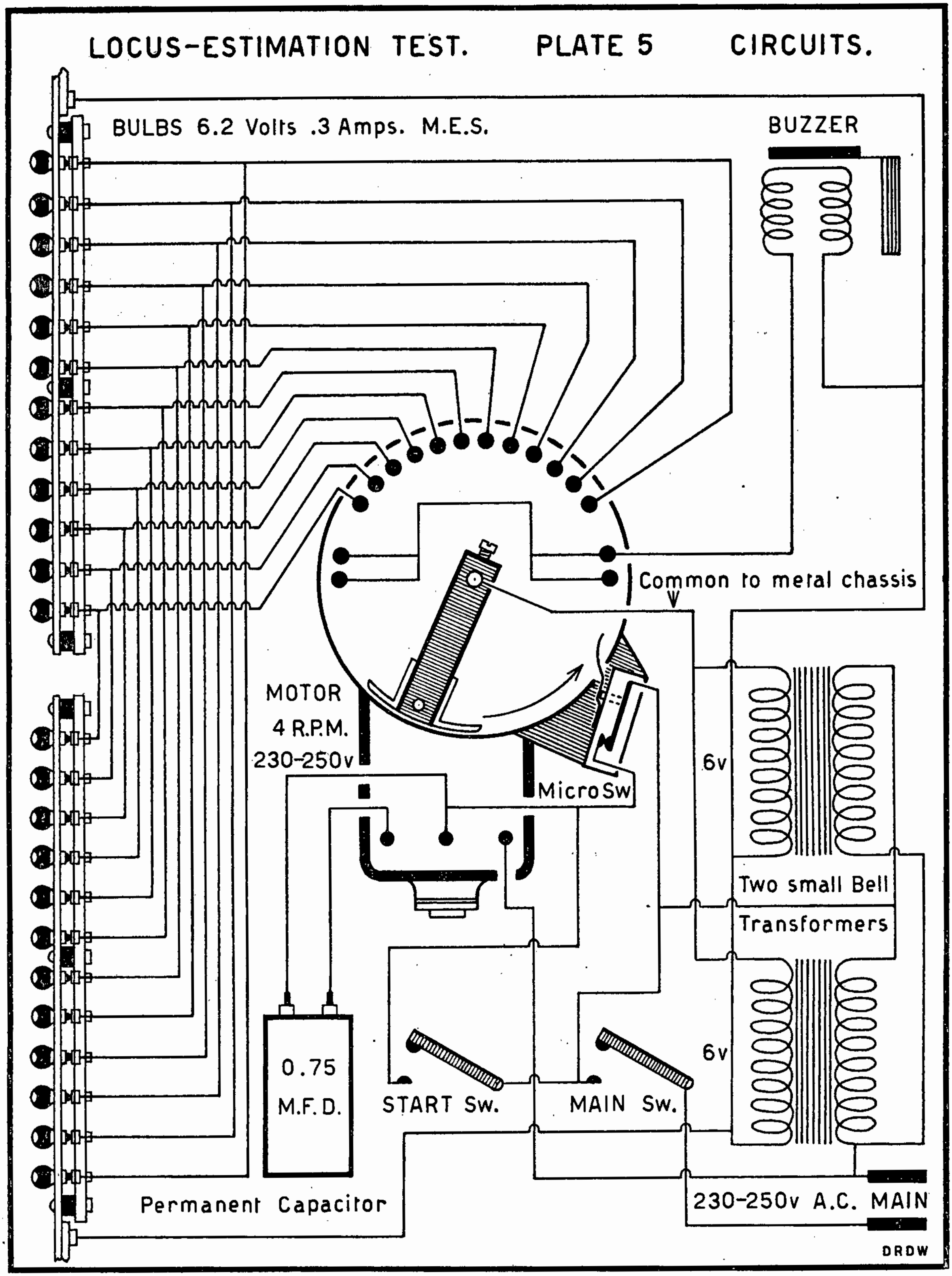
The Opal-glass screen, with numbers, letters, and graticule-lines scribed on the front or subject's side in matt-black lacquer, is supported by a rectangular wooden case. (Plates 1, 2 and 3.)

Midway inside the vertical ends (Plate 3), are located the pivot-mountings for shafts carrying the "light" bars and radius-rods by which these bars may be set at the different angles.

A back-board, removed in Plate 3 to show arrangement of components inside the case, is essential, because by reducing general illumination inside it ensures that subjects see only the light-spots, and no shadows or outlines of the bars which could divulge settings prematurely.


LOCUS-ESTIMATION TEST. PLATE 3 SETTING SYSTEM ETC.
For attachment to Hardboard back.
1/4 Holes
ARC STRIP M.S. 3/4 x 3/16
Space 1/2
Stubs of 2BA M.S. screws welded to edge.
Straight bar 18 1/8 perforated first then bent to an arc of Inside Radius 12 3/16.
Chamfers on holes
Rod turned down to 1/4 diam.
BSW 5/16
Stops M.S. Diam. 1/2 x 3/4 tapped for 4BA screw.
Round-end Peg 1/4 x 1/4
1 3/8
1 3/16
1/8
1 5/8
RADIUS ROD 1/2 diam. M.S. welded to collar 3/4 wide.
1/2
1 1/2
9 3/4
Knurled Knob
Drill out 1/4
Compression Spring 1 3/8
Retainer Frame for Glass Sheet 1 1/4 x 1/2
Inner edge of Main Front Frame rebated 1/2 x 1/8 to accommodate Glass.
Over-all dimensions of Glass Sheet 47 1/2 x 25
Interior of Case painted Matt Black
Frame for Back 1 1/4 x 3/4
5 3/4
1
1
36
11
18
8
OBA c/s
Rests Brass Angle 1x1x2
Transformers
2
Panel 8x6 1/2
1/4 Bakelite
Steel Bracket
Capacitor
13-Cord Cable to Lights.
Screwed from below.
3/4 Hard-wood.
15
8
To carry 'Light' Bar
3/4
3/16
1 11/16
15/64 Holes
3/4
5/16 x 3/16 deep
Recesses for Set-screws
M.S. Shaft
Diam. 1/2
Tap OBA
3/4
BSW 5/16
DRDW


LOCUS-ESTIMATION TEST. PLATE 4 LIGHTS & DISTRIBUTOR.
Brass Bar 5/8 x 3/16
Contact Strips Spring Brass 1/64 x 3/8
Bakelite Base 1 x 1/4
Steel Bar 3/4 x 3/16 welded to collar.
Holes tapped for M.E.S. bulbs 3/8 x 16 threads per inch.
13 1/8
17 1/4
1 3/4
3 1/2
4 BA
OBA
Aluminium Spacers 5/16 high x 1/2 Diam.
6BA
6 1/4
5 1/2
18 5/8
1 1/2
5/16 BSW.
Brass bearings brazed to Steel at A & B.
1 x 1/4 Steel Bar.
2 5/8
'KLAXON' MOTOR
Type: EM5CB1-M21
4 R.P.M.
Studs S1 & S2 on diameter.
Contact studs are heads of 2BA Brass screws turned blank and slightly convex. They project 1/32 above surface of disc and are recessed 1/32 below.
Extension pieces from Brass Angle 1 x 1
ROTOR
4 BA c/s
6BA c/h
DISC 6 Diam. Stationary on Motor Casing.
15/32
Clear area.
Lacquered Matt black.
Masking discs are stuck on bulb ends, lacquering completed, and discs then removed.
Discs are punched from adhesive masking tape.
Bakelite Sheet 1/4
3 3/8
2 5/8
7/16
2 BA
Brass Bar 3/4 square.
Recess for light spring
Micro Switch
1 9/16
Brass Sheet 1/16
Tap 4BA
2 1/2
Mount for Micro
Spacers Brass 1/2 diam. drilled and tapped 4BA.
MOTOR CASING
Brass Round
Terminal Nut
1 1/4
Bored out
Brass Brush
15/16
1 3/16
1 1/2
Set-screw grips on bearing case.
DRDW


LOCUS-ESTIMATION TEST. PLATE 5 CIRCUITS.
BULBS 6.2 Volts .3 Amps. M.E.S.
BUZZER
Common to metal chassis
MOTOR
4 R.P.M.
230-250v
Micro Sw
6v
Two small Bell
Transformers
6v
0.75
M.F.D.
START Sw.
MAIN Sw.
Permanent Capacitor
230-250v A.C. MAIN
DRDW

Both setting-systems, each consisting of its radius-rod with spring-loaded knob and perforated arc-strip (Plate 3) are situated outside on the back-board (Plate 2) and the perforations numbered legibly for convenient identification by the observer in making the settings.

In compiling a stock-list of the appropriate paired settings on X and Z for each of the fifty presentations and also the three examples, procedure was as follows:

Two points were marked on the main front frame so that they corresponded to theoretical extensions of the shafts carrying the "light' bars. With these points serving as pivots for two lengths of string, and the first and last bulb on each bar kept alight as indicators for aligning the bars with the strings, only such combinations of radius-rod settings were selected which gave in each case an intersection point of the loci well within some square on the screen.

Settings that produced intersection-points on the boundaries of squares were not accepted for the list.

It will be noted in Plate 4 that only a small round area at the end of each bulb is used for illumination. This is to obtain concentrated spots on the screen and also prevent superfluous lighting inside the case.

Electrical Circuits.

As shown in Plate 5, two small bell-transformers of the ordinary domestic pattern supply 6-volt current for the bulbs and buzzer. However, a single transformer of a larger type would serve as well.

Corresponding bulbs on each bar are wired in parallel and the pairs are energised through single contact-studs on the distributor. Synchronisation is thus positively maintained.

A micro-switch of the normally closed kind is arranged to switch the motor off automatically after each revolution. For starting purposes there is a manual switch in a loop by-passing the micro-switch.

The sound of the buzzer just before a presentation, in addition to warning subjects, informs the observer that the distributor-rotor has cleared the micro-switch, and the manual switch marked "Start" must be returned to its "Off" position. If this is not done the motor will not switch off automatically when the rotor engages the micro-switch again.

Summary and Conclusions.

(1) An apparatus-test designed primarily to assist in the selection of radar-screen observers for interception-control work, is described and discussed.

(2) In a preliminary study on a sample of 21 it yielded a correlation of .4 with the operational criterion.

(3) The functioning principle should be applicable to other laboratory equipment in the field of visual perception where an easily controlled apparent movement is required.

(4) It must be stressed, however, that this adaptation of the Phi-phenomenon does not provide a *perfect* succedaneum for actual continuous movement.

On the apparatus described, with the time intervals approximately .4 of a second and the bulbs spaced one inch apart, some residual saltation in stimulus progression is inevitable. If it is thought necessary to reduce this, then closer spacing of the bulbs or a higher rotor speed, or both, should effect an improvement.

Acknowledgments.

Mr. H. J. Groen took the photograph.

The instructions for the subjects were drawn up in collaboration with Mr. D. J. M. Vorster and Mr. J. Tromp.

The test was administered by members of the N.I.P.R. testing staff.

Mr. J. Tromp, who supervised the research on ground-control selection, kindly made test results available to the author.

GOLDFIELDS

TABLE I

Intercorrelations between tests
(Decimal points omitted)

N = 21	1	2	3	4	5	6
1 Mental Alertness						
2 Progressive Matrices	53					
3 Angles, Bearings and Directions	60	45				
4 Tables and Graphs	60	60	61			
5 Ships	63	55	57	46		
6 Locus-estimation	48	32_x	53	38_x	32_x	
Mean r	57	49	55	53	51	41

"x" implies that significance does not reach the 5% level

TABLE 2

Correlations between tests and occupational criterion

N = 21	r	p
Mental Alertness	0.59	0.01
Progressive Matrices	0.61	0.01
Tables and Graphs	0.58	0.01
Angles, Bearings and Directions	0.43	0.05
Locus-estimation	0.40	0.07
Ships	0.32	> 0.10
Personality Assessment	0.26	> 0.10

Reprinted from the Journal of the National Institute for Personnel Research.

J. NAT. INST. PERSONNEL RES., 1958, 7, 125—127

Technical Notes and Communications

A CANCELLATION TIMER

D. R. DE WET

(Received 5th April, 1957)

Many Cancellation Tests such as the Woodworth-Wells "Letters and Numbers" and others using grouped dots or geometrical symbols are most conveniently administered as group tests to a large number of subjects at once, time being a constant and performance measured in terms of quantity of matter handled by each subject.

For certain types of investigation, however, such as those in which fluctuations in performance have to be measured, it has been found necessary to administer cancellation tests to the subjects individually so as to secure a separate time measure for each line of stimuli.

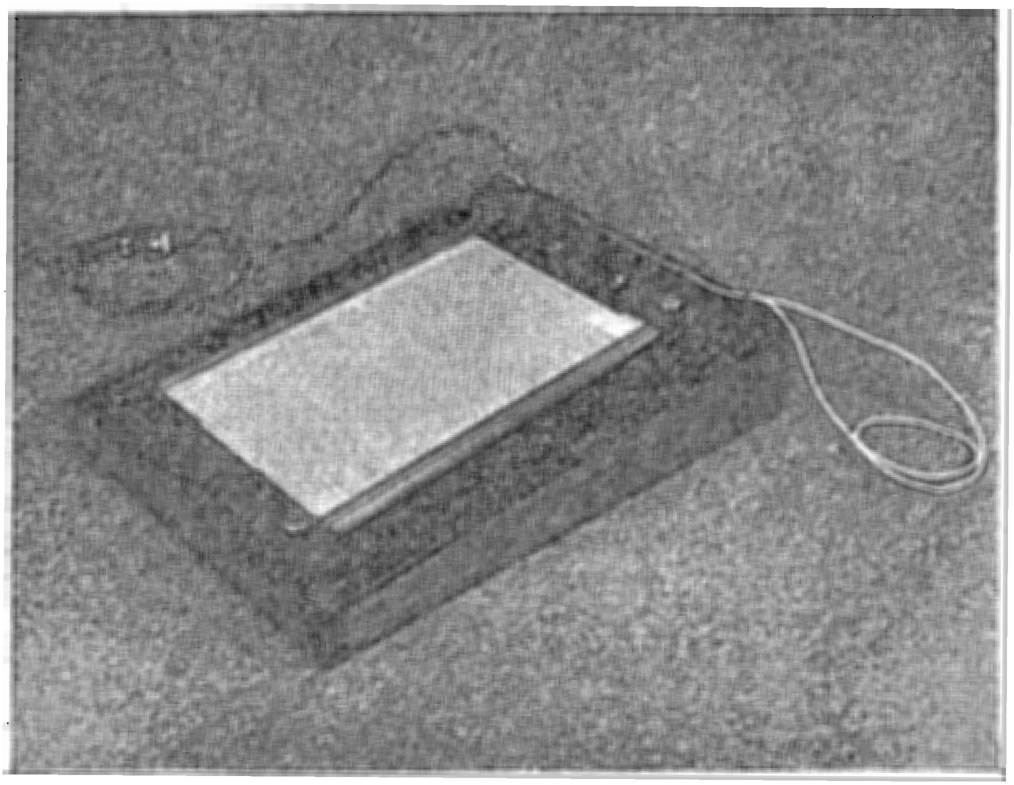
The method whereby the observer uses a stop-watch has been found unsatisfactory on account of the difficulty of judging the precise moment when the subject is satisfied in his own mind that he has completed the requisite cancellations on a line of stimuli.

Some other means, enabling the subject to time *himself* on each line, is necessary. But this should be of such a simple nature that it entails the minimum of effort distracting to his cancellation task.

General Description.

This timer device is an attempt to fulfil both the above requirements and appears on test to be successful. The additional task imposed on the subject is very slight. He uses a metal propelling pencil and touches it against metal strips bounding the ends of the lines. It is a simple routine which presents no difficulty to normal subjects.

The distance between the two strips is adjustable for different sizes of stimulus sheets, enabling the strips to be set up close to the ends of the lines and clamped in position. The apparatus consists of two table units :—



(a) A control-box housing the relays, buzzer and transformer and forming a mounting for the clocks.

(2) A lectern with a hinged frame to hold the stimulus sheet (Figure 1 and 3).

The top and bottom members of the frame are of non-conducting material such as bakelite and the side strips of metal. Recessed into the inner side of the lower bakelite member is a short metal strip marked "Stop".

Recording.

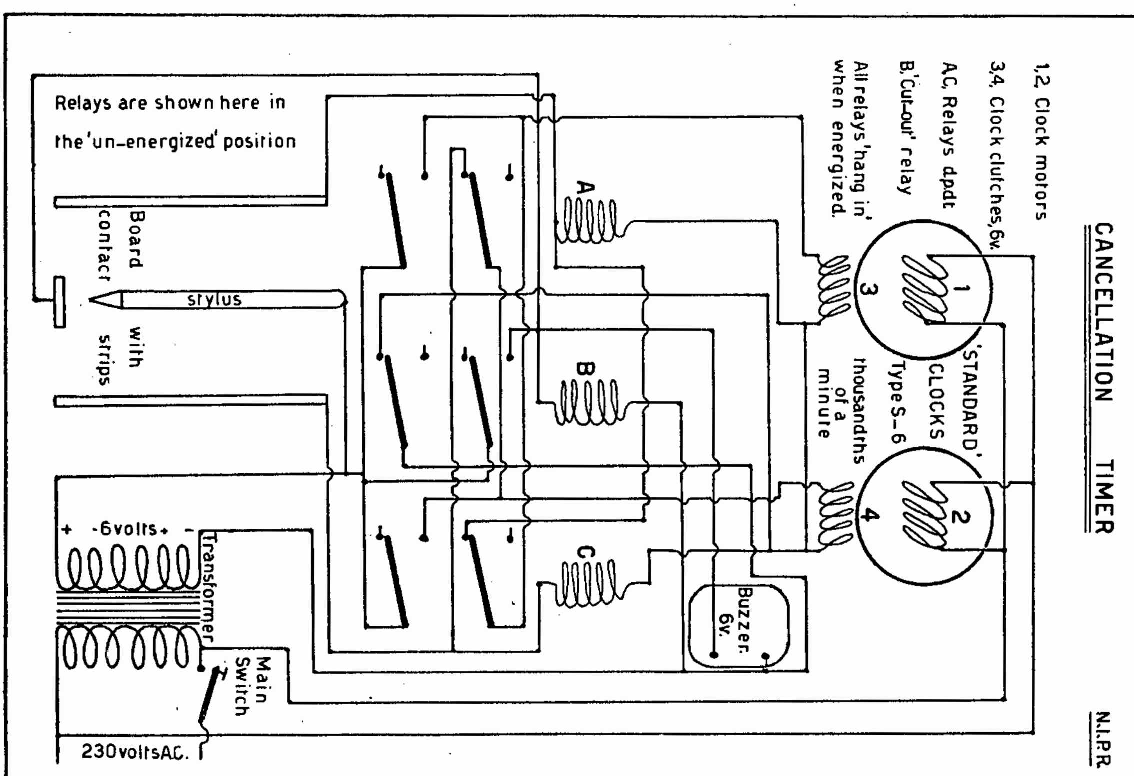
Recording of the subject's time period in the cancellation of the stimuli along the lines is done on two 'Standard' clocks (Type S-6, 1/1000ths of a minute), alternately for each successive line. While the one clock is running the observer takes down a reading from the other stationary clock.

On starting cancellation along the first line, the subject touches the left-hand metal strip with the metal side of his stylus. This brings clock No. 1 into operation. At the end of the line the right-hand metal strip is touched in a similar way; this stops clock No. 1 and starts clock No. 2 which records while he cancels back along the second line. On contacting the left-hand metal strip at the end of the second line, clock No. 2 is stopped and clock No. 1 started again to record cancellation along the third line, and so on alternately. When he has completed cancellation on the last line of the page, the short metal strip marked "Stop" on the lower frame member is touched with the stylus, which switches off both clocks and sounds a buzzer to apprise the observer that a trial is over and the main switch may be turned off.

Electrical Circuit.

The 'cut-out' system and the alternate transfer of recording to the clocks, utilize the principle of the 'hanging-in' relay. Referring to the wiring diagram (Figure 2), it will be seen that three separate double-pole-double-throw relays are employed, one each relating to a clock and one being the 'cut-out'.

A transformer supplies 6 volts A.C. to the clock clutches, relay solenoids and buzzer. The clock motors work directly off 250 volts A.C. When the main switch is put on the clock motors start and the transformer is potentialized but all three relays stay in the unenergised position and likewise the clock clutches.



With the 'cut-out' relay B in this position there is continuity from the one 6 volt transformer pole forming the common line to the two clock clutches and the solenoids of the two 'clock' relays A and C. The common line from this 6 volt pole to the solenoid of the 'cut-out' relay B and the buzzer is direct from the transformer and always has continuity, irrespective of whether the breaker points of relay B are closed or open.

The other 6 volt pole on the transformer is coupled to the stylus and one 'arm' contact on each relay, the function of which contact in each case is a dual one. When a relay is energised, the 'arm' contact transfers power to its own solenoid, causing it to 'hang in' and at the same time transfers power, in the case of relays A and C to their respective clock clutches, and in the case of relay B, to the buzzer.

Now in each relay where one 'arm' section makes contact to the relay solenoid and one outside solenoid (clutch or buzzer), the other 'arm' section breaks a line of continuity.

In relay A, the break cuts off the power to the 'hanging-in' system of relay C. In relay C the break cuts off the power to the 'hanging-in' system of relay A. In relay B the break cuts off the power to all 6 volt solenoids except the solenoid of relay B and the buzzer.

The relays A and C thus work reciprocally.

Contact between the stylus and the left-hand metal strip energises relay A and causes it to remain 'hanging in' when the stylus contact is broken. Also via this 'hanging-in' circuit, power is fed through to clock clutch 3.

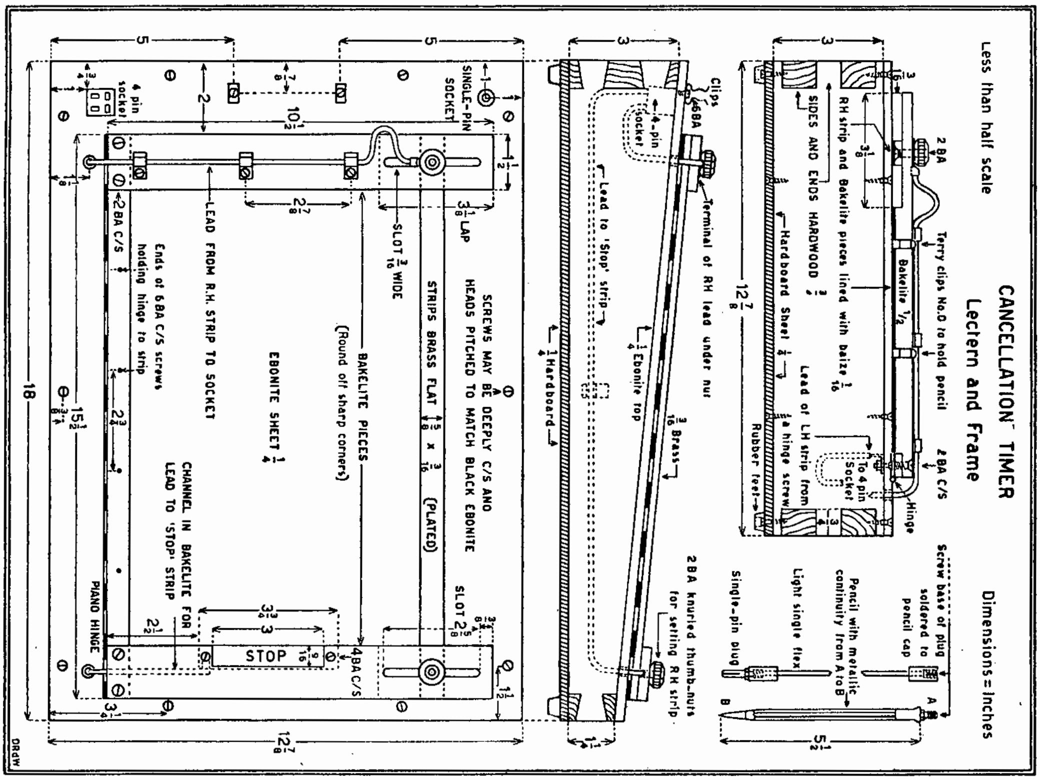
Contact between the stylus and the right-hand metal strip energises relay C which 'hangs in', breaks the 'hanging-in' circuit of relay A, de-energises clock clutch 3 and energises clock clutch 4. When the left-hand strip is contacted again, procedure is vice versa.

When the 'Stop' strip is contacted, relay B comes into operation, 'hangs in', sounds the buzzer and severs the common to the clock clutches and the solenoids of relays A and C.

The 'hanging-in' circuit of relay B is broken when the main switch is turned off.

Acknowledgement.

The apparatus was photographed by Mr. J. B. Kirstein.



GOLDFIELDS

Reprinted from "Psychologia Africana," Volume 11, 1965.

SOME ASPECTS OF PERFORMANCE IN A CANCELLATION TEST

D. R. DE WET

A Dots-cancellation test, purporting to measure 'sustained attention', was administered individually on a special timing device to a group of air-pilot candidates. Material and procedure are described. Results are discussed in relation to internal consistency, validity, correlation with other tests (intellectual, perceptual and sensory-motor), the kinds of error made, the effect of practice and the structure of the performance curves obtained with sub-groups of subjects who were respectively, good, middling and poor at the initial stage of the test.

Cancellation provides a simple perceptual-motor activity that has been found convenient by Piéron (14), Bugnion (3), Wiersma (18) and others for the measurement of 'voluntary attention', which Bugnion (3) regarded as a sustained discriminative task where the satisfaction is deferred, as distinct from 'spontaneous attention' where the satisfaction is immediate. If this distinction is acceptable, it would be better to substitute the term 'sustained' or 'continuous' for 'voluntary', because volition obviously plays an important part in both these functions of attention.

The motor element in cancellation involves no more than the repetitive activity of making pencil-strokes across printed items, and while there are undoubtedly individual differences in the time taken to perform this simple movement, its contribution to the total variance is probably less than that of the time taken in the perceptual activities of scanning the items and recognizing those that have to be marked. The task seems, therefore, to be rather more 'perceptual' than 'motor'.

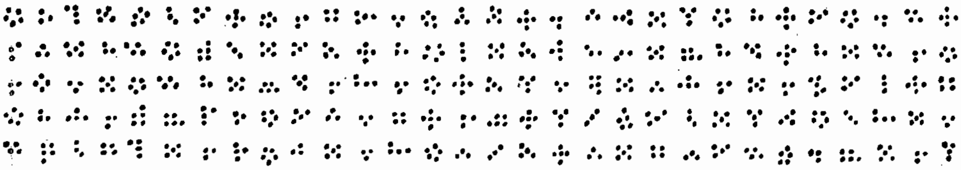
For general application, tests requiring the cancellation of abstract items such as directional signs, geometrical shapes, or groups of dots, are usually preferable to those where the items consist of standard symbols like letters or numbers, with which certain subjects are more familiar than others, as occurs very markedly in samples that are culturally heterogeneous.

A Dots-cancellation test, compiled by Dr. C. M. Mooney of Toronto, on the lines of a similar test used by Wiersma (18), who stated that he "prepared it from a model in an English periodical", was applied experimentally to a group of air-pilot candidates for the South African Air Force. The main objects were, to investigate: (a) Whether such a test of 'sustained attention' had any useful potentiality for the prediction of success in learning to fly; (b) Its internal consistence; (c) Its relationship to certain other tests; (d) The effects of practice on speed and accuracy; and (e) The relationship between speed and accuracy.

The present, preliminary study, is confined to forty of the candidates who were selected on the basis of other tests to undergo flying training.

Material and Procedure

The Dots-cancellation test-form consists of a sheet of paper 14″ × 9″ on which 1350 small groups of dots are printed in 45 parallel lines of 30 groups each, both vertical and horizontal spacing between the 'mid-points' of the groups being approximately a quarter of an inch. Each group is composed of three, four, or five dots; the sequence of the groups along the lines is irregular, and the patterns in which the dots are arranged, vary considerably. The first five lines, which are pretty representative of the whole test, appear thus:



It should be mentioned that there is a group of *six* dots, the only one of its kind in the whole test, at the beginning of line thirty-three; whether this was accidental or intentional on the part of the compiler, is not known.

In order to obtain accurate measures of the time taken to complete each line, the test was administered individually on the Cancellation timer (6), an apparatus designed specially for this purpose. The test-form is placed under an open rectangular frame of which the two long sides consist of straight metal strips adjusted so that the one is close to and parallel with the left-hand ends of the lines, and the other, the right-hand ends. At the lower edge of the frame, there is a short metal inset marked " Stop ".

Time, in thousandths of a minute, to complete successive lines, is recorded alternately on two clocks, which stop and start in turn as the subject touches the strips with the end of his pencil. While the one clock is running, the observer notes the reading on the other and sets the hands back to zero. After completion of the last line, contact between the pencil and the inset marked " Stop " switches off the current to the clocks and sounds a buzzer.

Instructions to the subject were as follows:

" On this paper there are various groups of dots. You must put a stroke through every group of five dots.

" Using this pencil in the hand with which you normally write, you touch it first against this metal strip (demonstrate) and then cancel each group of five dots along the line. As soon as you reach the end of the first line, you touch this metal strip (demonstrate) with the side of the pencil and then come back along the second line (demonstrate) putting a stroke through each group of five dots. As soon as you reach the end of this line, you touch this metal strip again (demonstrate) and then you go out again in the third line in the same way, return on the fourth, and so on all the way down. Each time you reach the end of the line you must touch the metal strip.

" There is one thing I want you to try to remember. When you reach the end of the last line, don't touch this strip as you have been doing, but touch this one marked ' Stop ' (demonstrate). Work as quickly as you can and try not to miss out any of the ' five-dot ' groups or cancel any of the other groups ". (The subject is arranged so that he sits comfortably and has the lectern holding the test-form at the distance and angle he finds most suitable.)

" You may begin!"

Results and Discussion

Reliability

As indicated by the split-half correlation coefficients in Table 1, the internal consistency of the Time measure is high and that of the Error measure, although considerably lower, still acceptable. The two kinds of error, namely, omission of items that should have been cancelled and wrong cancellation were combined here, for neither kind occurred very often. It is to be expected that 'Error' should be less reliable than 'Time', because errors are comparatively rare events and therefore more subject to the vagaries of chance. (The ratio of error to correct response is about 10:1340.)

Table 1

Product-moment correlations between first and second half of the test

	Time	Errors
Uncorrected ..	0·93	0·61
Corrected ..	0·96	0·76

This is also the main reason, why, in Table 2, the distribution of 'Error' scores, either separate or combined, is of the Poisson rather than the normal type and the SD's of these scores (Table 3) are about as great as the means. In both the 'Time' and 'Error' distributions there is the inevitable irregularity characteristic of small samples.

Table 2

Frequency distributions of Time and Errors

TIME		ERRORS			
Minutes	N		Wrong Cancel.	Omit. Cancel.	Total Wr./Omit.
		0	1	1	1
8·3— 9·3	1	1— 3	25	13	3
9·4—10·4	6	4— 6	9	9	12
10·5—11·5	6	7— 9	3	10	3
11·6—12·6	12	10—12	1	3	8
12·7—13·7	5	13—15	0	0	8
13·8—14·8	7	16—18	0	0	0
14·9—15·9	2	19—21	0	2	0
16·0—17·0	0	22—24	1	2	0
17·1—18·1	0	25—27	0	0	3
18·2—19·2	1	28—30	0	0	2
N =	40			40	

Referring to Table 3, it will be seen that errors of omission in cancelling five-dot groups occurred nearly twice as frequently as wrong cancellations of other groups. Behaviourally, wrong cancellation is the more 'serious' error, for omission means only incorrect perception, whereas wrong cancellation means incorrect perception confirmed by the closure of a misplaced motor response. In life generally, there is the analogy that sins of omission are usually less serious than sins of commission, but occur more often.

As could be expected, wrong cancellation was done more frequently on four- than three-dot groups. There is a small positive correlation between the two kinds of error which on the present sample does not reach the five percent level of significance.

Table 3

Means of the two measures of error

Error	Mean	S.D.	P of Diff.
Omitted Cancellations ..	6·8	5·6	<0·003
Wrong Cancellations ..	3·8	3·9	
Product-moment correlation between the two measures of error: 0·21 (Not significant)			

Improvement in Performance

Comparative improvement in speed and accuracy of work, as reflected by the differences between the respective means for the first ten lines and the last ten lines of the test, is shown in Table 4. Improvement in speed is very significant, and the graph (Figure 1) is a fairly typical learning curve, steep during the early stages of the task and then progressively more gentle. The curve would have been smoother, were it not for the fact that the number of five-dot clusters to be cancelled in each line varied from eight to thirteen, the average for the forty-five lines being 10·6. While it was not thought worthwhile to compute significance levels, there seems to be a positive relationship between fluctuations in the curve and the number of items to be cancelled. The five main upward fluctuations at lines 7, 15, 21, 32 and 40, relate respectively, to 11, 12, 9, 10 and 8, items; and the five main downward fluctuations at lines 8, 16, 22, 25 and 34, to 12, 13, 11, 12 and 11, items. Three of the former group involve less than the average number of items, and all the latter, more than the average. In a cancellation test of this kind, variation in the number of cancellable items per line has the advantage of preventing the subject from checking his responses by counting to ensure that the fixed number has been made in each line, but otherwise it is undesirable, as fluctuations in performance relating to more interesting factors like fatigue and vacillation of attention, become obscured.

Table 4

Improvement in speed and accuracy between the first and last ten lines of the test

Measure	Mean	S.D.	t of Diff.
Time (Mins.)			
First ten ..	3·17	0·52	15·52
Last ten ..	2·51	0·46	
Omit. Cancel.			
First ten ..	1·93	2·28	1·50
Last ten ..	1·40	1·57	
Wrong Cancel.			
First ten ..	0·58	0·59	1·07
Last ten ..	0·78	1·48	
Wr./Omit.			
First ten ..	2·48	2·48	0·72
Last ten ..	2·18	2·26	

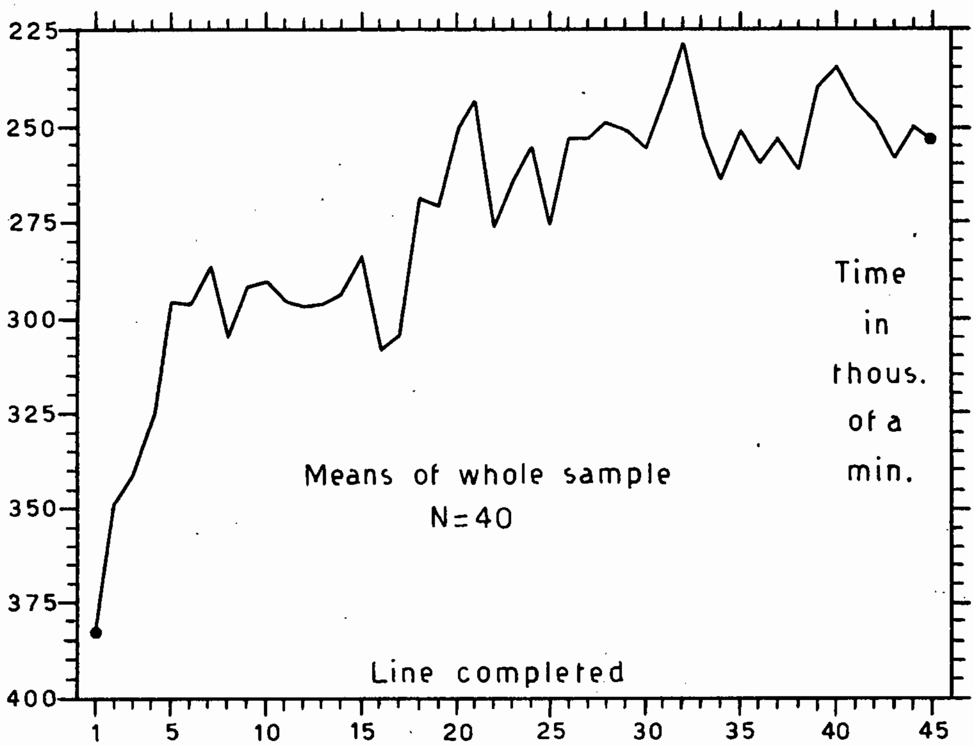


Figure 1. Average time in thousandths of a minute taken by the whole test sample to cancel the items in each of the forty-five lines (300/1,000 of a min. = 18 seconds).

The test sample as a whole showed no improvement in accuracy of performance. In neither omission, wrong cancellation, nor both combined (Table 4) is there any significant difference between the mean of the first ten lines and that of the last ten lines. A graphic representation of the frequency of the two kinds of error is given in Figure 2. The general preponderance of omissions over wrong cancellations is obvious, but at line thiry-three, on the ‘ wrong-cancellation ’ profile, there is an interesting anomaly. As mentioned previously, the first item in this line consisted of a cluster of *six* dots, the only one of its kind in the whole test. Although the instructions emphasized that only five-dot items should be cancelled, thiry-nine of the forty subjects cancelled this one. Evidently, the solitary subject who avoided the snare did not do so merely by good luck, for he was the *rara avis* (Table 2) who made no error at all, either of omission or wrong cancellation.

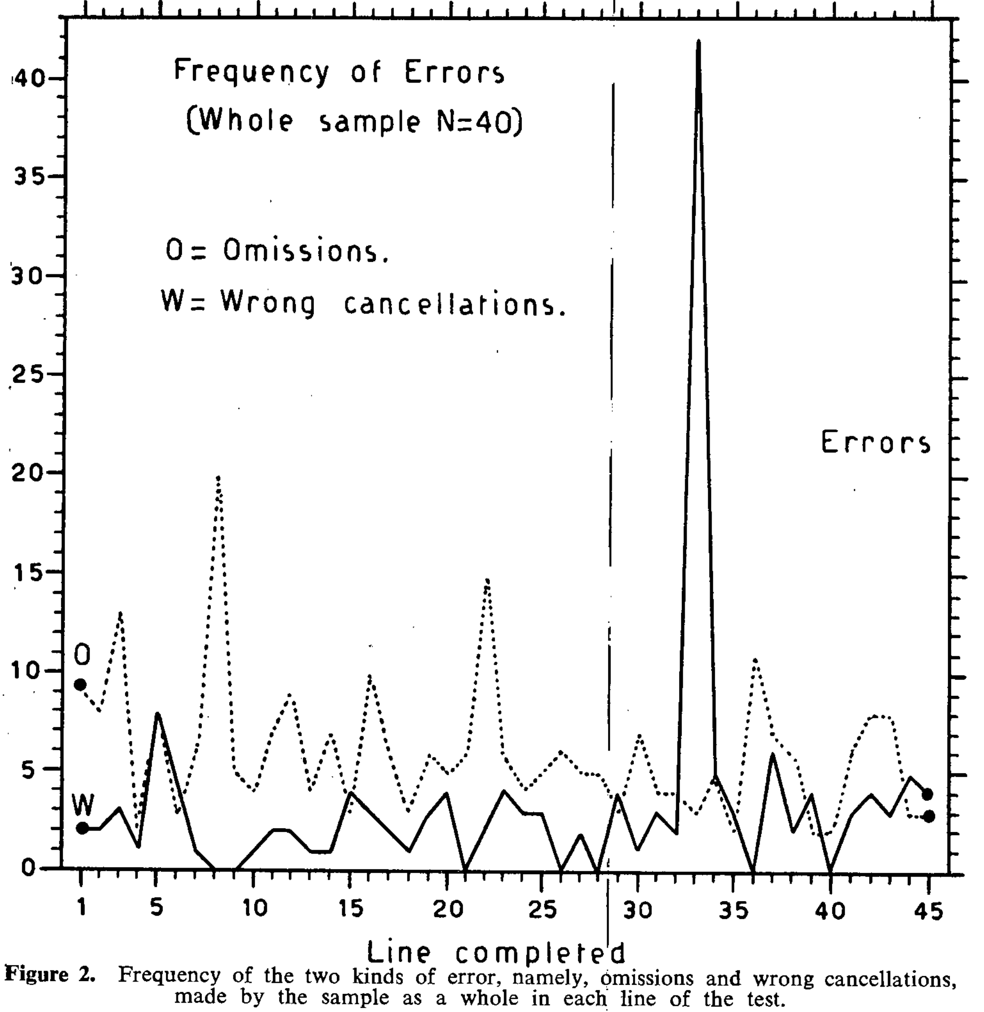


Figure 2. Frequency of the two kinds of error, namely, omissions and wrong cancellations, made by the sample as a whole in each line of the test.

It is likely that more than one influence contributed towards the deceptiveness of the six-dot item. Perhaps the most significant was the element of ‘ expectation ’. The possibility of there being an item of more than five dots was not mentioned in the instructions, and progress through the greater part of the test reinforced the notion that five was the maximum. Secondly, after a certain amount of practice, subjects would tend to given less attention to individual dots composing an item and more to the density and area of the overall *gestalt,* which afford quick indications of items that have lots of dots. In effect, the emphasis is transferred from the 'five-dot' criterion to the more general principle of reacting whenever there are "lots of dots".

Validity

As shown in Table 5, neither speed nor accuracy of performance in the Dots-cancellation test is related to success or failure in learning to pilot aircraft. The mean test scores of the eight subjects who were suspended because of poor flying ability are not significantly different from those of the thirty-two who succeeded on the course. There was no difference either, between the proportions of Successes and Failures who forgot to touch the inset marked "Stop" at the end of the test. Exactly one-quarter of either group omitted to carry out the instruction. Nor does rate of improvement in Dots-cancellation as reflected by the difference between performance in the first and second half of the test, relate to performance on the flying course. Sustained attention of the kind measured by this rather 'clerical' test is apparently unimportant in the highly practical vocation of the air-pilot.

Table 5

Dots-cancellation and Flying Training

Measure	Successes: 32		Suspensions: 8	
	Mean	S.D.	Mean	S.D.
Total Time	12·4	2·0	12·3	1·4
Omitted Cancellations ..	6·9	5·2	6·4	7·3
Wrong Cancellations ..	3·7	4·2	4·1	2·9
No significant difference between means				

Correlations with other tests

Correlations obtained between the Dots-cancellation Test and certain other tests, intellectual, perceptual, and sensory-motor, are given in Table 6. Speed of cancellation is moderately associated with both speed of work and error-proneness on Moede-type Two-hand Co-ordination, and speed of work, only, on Serial Discrimination (finger responses on a keyboard to visual stimuli) and Handlebars Co-ordination. There appear to be slightly more substantial relationships between errors in cancellation (both kinds combined) and Mental Alertness, Progressive Matrices, and errors in Serial Discrimination. The correlations with the two intellectual tests may be partly attributable to the fact that these tests and Dots-cancellation have a common 'pencil-and-paper' element. (In an early investigation, Bousfield (2) obtained low positive correlations between mental ability and performance in duplicating simple geometrical figures. Here there was also a common 'pencil-and-paper' element.)

There is no relationship between performance on Dots-cancellation and measures of speed of perception and span of attention taken on a tachistoscope (7); the difference may be due to the fact that the former requires sustained attention and the two latter, not.

On the present sample, the small negative correlation (0·26) between speed and accuracy of cancellation does not reach the five per cent level of significance. It is probable, that to a slight extent, fast working is associated with less accuracy and slow working with more.

Table 6

Product-moment correlations between Dots-cancellation and other Tests (N: 40)
(Decimal points have been omitted)

Other Tests	Dots-cancel. Time	Dots-cancel. Error
Mental Alertness	16	−43 S
Progressive Matrices (Raven's)	−01	−39 S
Foveal Flicker Fusion	08	−07
Speed of Perception	08	04
Span of Attention	−08	06
Two-hand Co-ord., Moede Type (Error)	−32 S	07
Two-hand Co-ord., Moede Type (Time)	32 S	00
Hand-foot Reaction (Speed of response)	−04	−07
Hand-foot Reaction (Total Error)	−16	16
Serial Discrimination (Time)	33 S	16
Serial Discrimination (Error)	−18	37 S
Handlebars Co-ord. (Time)	35 S	−01
Handlebars Co-ord. (Error)	04	15

Product-moment correlation between Time and Error on Dots-cancellation test: −0·26 (Not significant)

" S " implies significance at better than the 5% level

Parallelism in performance curves

Some indication of the relationship between initial ability at the cancellation task and subsequent performance, is afforded by the graphs in Figures 3 and 4. On both " Time " and " Error " measures there is a tendency towards parallelism between the work curves for subjects of high, median and low initial ability. In agreement with the findings of Ehinger (8) and others on various sensory-motor tests, subjects with low initial ability improve more markedly in relation to their own starting level than those with better initial ability. This applies particularly to speed of work in Figure 3, where all three groups show a degree of improvement between start and finish that is significant at well above the five per cent level. With regard to the combined " Error " measure in Figure 4, although the sample as a whole did not show any significant improvement in accuracy of work, it would appear that this relates mainly to those subjects whose initial standard of accuracy was median or good. Those who were initially very inaccurate do show some improvement, significant at the five per cent level, between their mean scores for the first and last ten lines of the test.

The obvious parallelism between the curves in both Figures 3 and 4 indicates that, practice being equal, the high starters tend to retain their ascendancy over the average and low starters, and likewise, the average over the low starters. Similar results have been reported by investigators using various other tests involving mainly perceptual and motor elements, namely: Reynolds and Adams (15), One-hand regular pursuit; Cieutat and Noble (4), Two-hand irregular pursuit; Mukherjee (11), Two-hand self-paced; Hudson, Mokoatle and Mbau (10), the " Chopsticks " test (5) and the factory-operation of " Nutting "; Abbey (1), Control-display relationships involving lateral and fore-and-aft movements of a lever on an apparatus known as the Toronto Complex co-ordinator, which has been described by Shephard (16); Phillips and Summers (13), Bowling; and Farmer (9), Simple reaction-time, Ordinary choice-reaction, and Complex choice-reaction where the response keys were not directly identifiable and the subject had accordingly to work by reference to a chart or by memory.

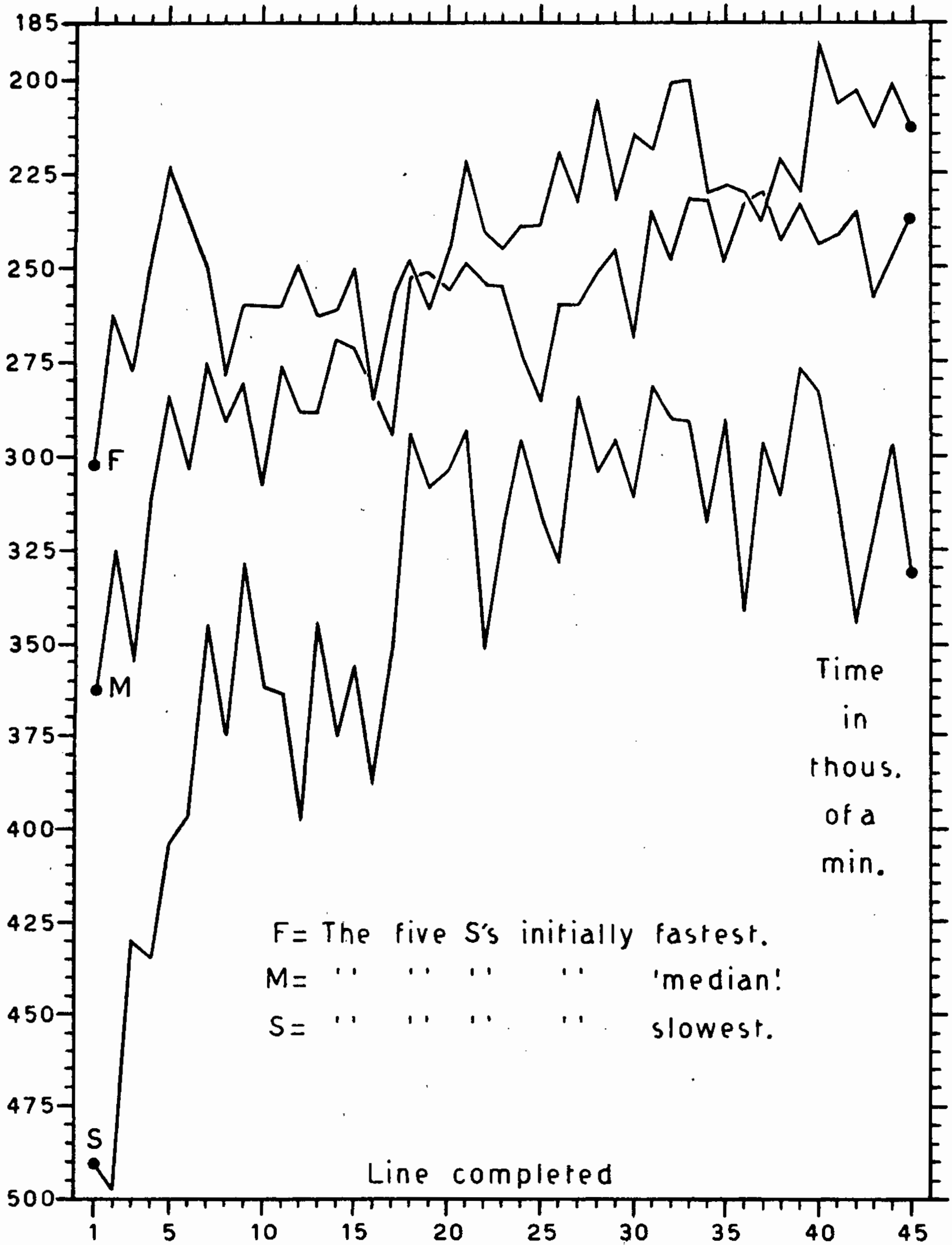


Figure 3. Average time in thousandths of a minute taken to cancel the items in each line by the three respective sub-groups who were initially fastest, median and slowest.

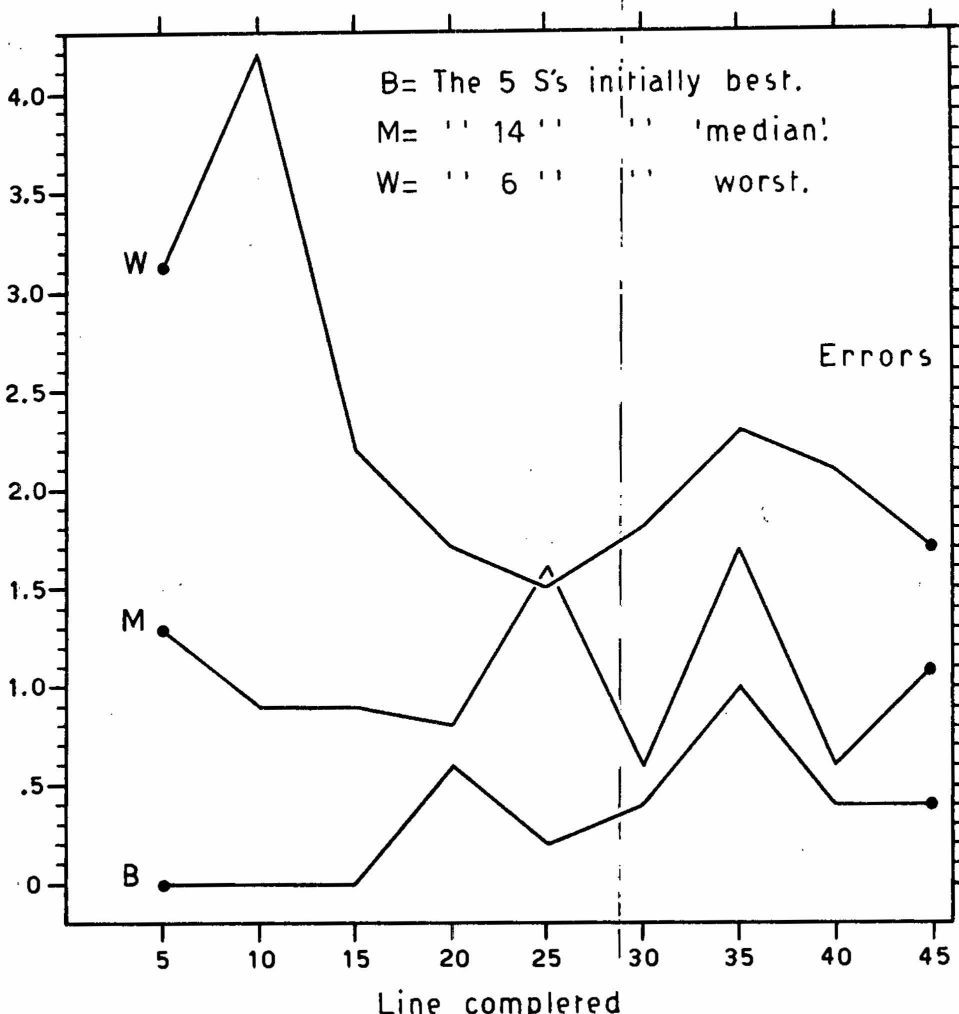


Figure 4. Average number of errors (omissions and/or wrong cancellations) made in each consecutive section of five lines by the three respective sub-groups who were initially best, median and worst.

Farmer (9), who seems to have been the first to observe this general relationship between initial level of ability and subsequent performance, concluded "... that curves of motor performance whether derived from psychological tests or industrial work, tend on the whole to be parallel". He also mentioned that this has considerable industrial importance in that improvements towards simplification in factory tasks will affect all the operatives similarly, though not necessarily to the same extent.

From the standpoint of personnel selection by aptitude tests, this has the very significant implication that comparatively short tests, provided, of course, that their reliability is adequate, can furnish useful measures of ability, and there seems to be little practical advantage in the extended repetition of trials as recommended by Wulfeck (19), although this may well produce some increase in reliability.

While the term ' parallelism ' is a convenient one to describe the general loci followed by the performances of the initially good, middling and poor subjects, it must not be taken too literally, because there can be appreciable variation in the degree of parallelism between the sets of curves obtained on different tests. The findings of Hudson, Mokoatle and Mbau (10) on the Chopsticks test show that at the end of a 21-day learning period there was considerably greater differentiation between the high, average and low starters, than at the beginning, whereas on the Nutting machine the separation between the groups at the beginning and end was not significantly different. " Chopsticks " calls for a fairly high degree of fine co-ordination, provides a very wide range of skill measurement, and is a test which most subjects find intrinsically interesting. On the other hand, Nutting requires a comparatively low order of skill, is largely a matter of speed of movement, and is generally regarded as a rather boring task. The more attractive ' challenge ' provided by Chopsticks may well have had some effect in making the average starters increase their initial lead over the below average, and similarly, the good starters, over the average. Knowledge of results as a reinforcement of the incentive should also be taken into account. In certain practical tests like Chopsticks, errors are apparent to the subject, but in Dots-cancellation there is no obvious indication to him that he has made an error.

Convergence

Some tendency towards gradual convergence is more usual with performance curves, as shown in Figures 3 and 4, and also in the graphs of Mukherjee (11) and Phillips and Summers (13). (This is also reflected in a decrease in SD for the whole sample, between initial and final performance, as in Table 4.) After extensive practice the differences in proficiency levels between the groups are substantially less than at the beginning of the task, although the hierarchy of the groups remains the same.

Another respect in which the curves of Hudson, Mokoatle and Mbau (10) differ from those of the writer and others, is that the initially poor performers did not show more substantial improvement than those who were initially better. This may be concerned with the nature of the test sample, which consisted of African factory operatives whose standard of formal education was very low, whereas the present sample and those of Abbey (1) and Wulfeck (19) consisted of educated Europeans, and that of Mukherjee (11), of educated Asians. Investigations on the abilities of Africans by Vernon (17), Murray (12) and others, have disclosed a certain simplicity of factor structure which suggests that African general ability is far more intimately bound up with a concrete orientation and the manipulation of physical things than is, what Vernon calls, " British g ". On this hypothesis, the same perceptual-motor task may involve considerably more " intelligence " in one cultural group than another. Vernon (17; p 84) said of his first factor, that it " possibly represents general adaptability to the unfamiliar testing situation rather than the g obtained with educated whites ". If the poorer groups on Chopsticks and Nutting were made up mainly of those subjects who were also the least intelligent (as reflected by their initial performance on these tests), or who adapted badly to the ' unfamiliar testing situations ', it is understandable that they failed to show a more substantial improvement than the groups above them.

Apart from cultural influences, the degree of convergence in the performance curves of subjects at different levels would seem to depend mainly on the scope for the exercise of skill provided by the task and the extent of practice. Plenty of practice at a task affording a rather low range of difficulty, very markedly reduces the differences between the levels. An example of this is to be found in the results of Wulfeck (19) whose subjects did ten trials daily for ten days on a pursuit two-hand co-ordination test in which two handles were rotated to keep a stud over a small target moving along an irregular path. The curve of the five best subjects, and that of the five worst (rated on a preliminary series of trials) ran remarkably parallel until the fifth day, and thereafter rapidly converged on to the same plateau. In such cases, practice apparently evens out individual differences in initial ability because the low ceiling of difficulty in the task prevents any further discrimination between superior and inferior performers. This result does not invalidate the concept of a general parallelism in performance curves. As long as the activity affords scope for improvement, it is probable that the initially good performers will continue to keep ahead.

Summary and Conclusions

1. A Dots-cancellation test, providing a simple measure of "sustained attention" was administered individually to a group of pupil air-pilots on the Cancellation-timer apparatus (6).
2. The "Time" measure is more reliable than the "Error" measure, which has, however, acceptable reliability.
3. Errors of omission occurred nearly twice as often as the more "serious" errors of wrong cancellation.
4. All subjects showed a significant improvement in speed of work.
5. The test sample, as a whole, showed no improvement in accuracy of performance, but those subjects who were initially very inaccurate, did show some improvement, significant at the five per cent level.
6. Thirty-nine of the forty subjects made the mistake of cancelling an odd cluster of six dots, probably because of the element of "expectation" and the fact that this cluster occurred at an advanced stage of the test after certain associations between *gestalt* and cancellability had become well established.
7. Neither speed nor accuracy of performance in the Dots-cancellation test is related to success or failure in learning to fly.
8. Speed of cancellation is moderately associated with speed in a test of serial discrimination, speed in two tests of co-ordination, and also error-proneness in one of the latter. Slightly more substantial correlations were obtained on the cancellation error measure with achievement in two intellectual tests and error in serial discrimination.
9. To a barely significant extent fast working is accompanied by less accuracy and slow working, by more.
10. In both speed and accuracy of cancellation there is a tendency towards parallelism between the performance curves of subjects who are initially good, median and poor. Initially poor performers show the most marked improvement in relation to their own starting level. Relevant findings of other investigators on other perceptual-motor tests are discussed.

ACKNOWLEDGMENTS

I am grateful to Miss M. C. Cooper and Miss M. E. Jones for their assistance in administering the test, and to Dr. H. Reuning for constructive criticism of the manuscript.

REFERENCES

1. ABBEY, D. S. Ability and performance on a complex perceptual-motor task. *Percept. Mot. Skills*, 1960, 11, 55–56.
2. BOUSFIELD, W. A. A study of motor skill in the free hand duplication of geometrical figures. *J. appl. Psychol.*, 1930, 14, 478–485.
3. BUGNION, B. Le test de barrage. *Arch. Psychol., Genève*, 1935, 25, 77–97.
4. CIEUTAT, V. J. *and* C. E. NOBLE. Ability versus practice in two-hand co-ordination. *Percept. Mot. Skills*, 1958, 8, p. 226.
5. DE WET, D. R. Two tests of implement manipulation. *J. Nat. Inst. Personnel Res.*, 1957, 7, 75–79.
6. DE WET, D. R. A cancellation timer. *J. Nat. Inst. Personnel Res.*, 1958, 7, 125–127.
7. DE WET, D. R. Measures of speed of perception and span of attention on air-pilot candidates. *Psychol. Afr.*, 1964, 10, 206–218.
8. EHINGER, G. Recherches sur le développement de l'habileté manuelle par la pratique d'un métier manuel. *Arch. Psychol., Genève*, 1927, 20, 299–317.
9. FARMER, E. Parallelism in curves of motor performance. *Brit. J. Psychol.*, 1927, 17, 335–342.
10. HUDSON, W., B. MOKOATLE *and* G. G. MBAU. The influence of training and practice in the test and work performance of a small sample of African workers. *J. Nat. Inst. Personnel Res.*, 1958, 7, 88–94.
11. MUKHERJEE, B. N. Learning efficiency in a psychomotor test as a function of initial skill. *Engineering and Industrial Psychology*, 1959, 1, 138–142.
12. MURRAY, C. O. *The Structure of African Intelligence. A Factorial Study of the Abilities of Africans.* M.A. thesis. Durban, University of Natal, 1956.
13. PHILLIPS, Marjorie *and* D. SUMMERS. Bowling norms and learning curves for college women. *Res. Quart. Amer. Ass. Hlth. phys. Educ.*, 1950, 21, 377–385.
14. PIÉRON, Mme. H. Etalonnage d'un test d'attention. Le test de Toulouse et Piéron. *Bulletin de l'Institut National d'Etude du Travail et d'Orientation professionelle*, 1929, 1, 105–112.
15. REYNOLDS, B. *and* J. A. ADAMS. Psychomotor performance as a function of initial level of ability. *Amer. J. Psychol.*, 1954, 67, 268–277.
16. SHEPHARD, A. H. The Toronto complex co-ordinator. *Canad. J. Psychol.*, 1955, 9, 7–14.
17. VERNON, P. E. The structure of practical abilities. *Occup. Psychol.*, 1949, 13, no. 2, 81–96
18. WIERSMA, E. D. *Lectures in Psychiatry.* London, H. K. Lewis, 1932.
19. WULFECK, W. H. Learning the two-hand co-ordination test. *J. appl. Psychol.*, 1942, 26, 41–49.

Mr. D. R. de Wet,
National Institute for Personnel Research,
P.O. Box 10319,
Johannesburg,
Republic of South Africa.

Manuscript received: 4th January 1965.
Revised manuscript received: 12th February 1965.

PRINTED BY CAPE & TRANSVAAL PRINTERS LIMITED, JOHANNESBURG

Reprinted from "Psychologia Africana," Volume 10, 1963.

A DUAL-PURPOSE "FALL" TACHISTOSCOPE

D. R. DE WET

This instrument was designed for the presentation of either moving or stationary visual stimuli for short periods. It is fairly compact, of rugged construction, and having basically one moving part is somewhat more accurate and mechanically dependable than certain early types of "fall" device with more involved linkage systems. Its main limitation, in common with other mechanical tachistoscopes, is that it has too long a minimum exposure-time for certain work requiring the liminal exposure of stationary stimuli. Mundy-Castle (3) experienced this when, in the absence of a faster instrument, he was compelled to use the present one. However, the restriction does not occur in the case of moving stimuli, and for many practical purposes liminal exposures can be obtained by utilizing a moving stimulus instead, as was done by Biesheuvel and Mundy-Castle (1).

Description (Plates 1 to 5)

A metal beam of rectangular cross-section pivots vertically in two ball-races mounted in a frame held by two knurled nuts to the back member of a hardwood chassis. At the front end of the beam a segment of sheet aluminium with a rectangular opening at the middle is rigidly attached (Plates 2 and 4), and at the tail-end, there is a counterpoise-weight and a smaller trimming-weight, both adjustable and fitted with set-screws (Plates 1, 2 and 5). A light cursor-weight slides along the other arm of the beam between the pivot and the segment, and by means of its set-screw can be fixed opposite linear graduations engraved on the side of the arm (Plates 2 and 4).

In the front of the chassis there is a window about the same size as the rectangular opening in the segment and directly opposite it when the beam is horizontal (Plates 1 and 3). Close behind the window, a flat metal shield with a small finger-knob, is pivoted, and when the knob is depressed the shield rises and obscures the window (Plates 2 and 3). A rectangular frame with grooves along three sides to accommodate light cards ($3\frac{5}{8}'' \times 2\frac{3}{8}''$) is held by two screws and spacers a short distance behind the window and in alignment with it (Plates 2 and 5). Metal strips spaced around three sides of the opening in the segment provide a similar card-holder here (Plates 2 and 4).

Inside the top of the chassis, a flat metal bar covered with felt and adjustable vertically to a slight extent by means of a nut on each of the supporting screws at its ends, serves as a buffer for the upper edge of the segment when the long arm of the beam is raised (Plates 2 and 3).

Extending up at an angle from the base-board of the chassis, is a metal bar ending in a felt-padded fork and fitted with a small spring-loaded finger-catch (Plates 2 and 5).

The beam-arm and segment are held at their topmost position by a spring-loaded catch pivoting on a plate bracketed near the top of the longer side-member of the chassis (Plates 2, 3 and 5), and release is effected by a pneumatic piston located above the catch (Plates 2 and 5) and actuated by pressure on a rubber bulb. When the beam-arm is released, it carries the segment down between the back of the window-shield and the front of the stationary card-holder, comes against the felt pad in the fork of the stop, and is prevented from rebounding by the finger-catch which automatically locks in place over a short projection (Plates 2 and 4).

The purpose of having the frame which supports the beam-fulcrum held by knurled nuts to the chassis instead of having it permanently bolted, is to permit a simple lateral adjustment to ensure that the end of the beam is in proper alignment with the pneumatic catch. When the beam is cocked, its end should be in the middle of the fork in the small guide-piece situated behind the catch near the lower right-hand corner of the mounting-plate (Plate 5), so that it can fall freely when the catch is released. Extremes of atmospheric temperature or humidity, or the fact that the apparatus is resting on an uneven surface might produce sufficient twisting of the chassis to cause the beam-end to come against one or other of the sides of the guide-piece fork at the cocked position and impede free falling. Before starting a test series, particularly if the apparatus has not been in use for some time, it is advisable to check that the cocked beam-end is " free " in the middle of the fork. If necessary, the two knurled nuts are slackened, the beam-end moved over a trifle, and the nuts tightened again.

Procedure (Plate 1, lower figure)

In most applications of this apparatus, it is preferable for the subject to control the actual presentation of the stimuli himself. This helps to keep the expectation element reasonably constant and ensure that exposures are made only when he is ready for them.

He is seated before the window and his distance from it adjusted until he is at an optimum position for reading ordinary printed or typed matter on a card held in it.

Whether stationary or moving stimuli are presented, a test series may be done either by setting the cursor-weight back a certain distance towards the beam-pivot for each trial and so giving consecutively longer exposures of a particular card until the subject can identify the stimulus, or by retaining a constant speed with the cursor-weight fixed at one position and exposing consecutive cards each having an increased number of discrete stimuli to be identified.

When stationary stimuli are to be presented, the observer " cocks " the beam-arm at its top position, puts the first card in the holder fixed opposite the window, and sets the cursor-weight either at some standard pre-determined graduation mark for constant-speed exposure of consecutive cards, or at some other pre-determined mark near the end of the arm to give the shortest of the intended consecutive exposures of the same card.

The shield behind the window is then lowered, and the subject told to watch carefully and press the bulb, which he holds in his preferred hand. The segment falls and the stimulus is exposed while the open rectangular section passes it. The observer notes the subject's report on what he has seen, raises the shield to close the window, releases the beam-arm at the stop by depressing the finger-catch and re-cocks the arm.

If the same card is to be exposed again, the cursor-weight is moved back for a standard distance, the shield lowered, and the subject told to watch and press again. This is repeated until the cursor-weight is at a graduation mark where the arm falls just slowly enough for the subject to identify the stimulus correctly, and the reading is noted by the observer.

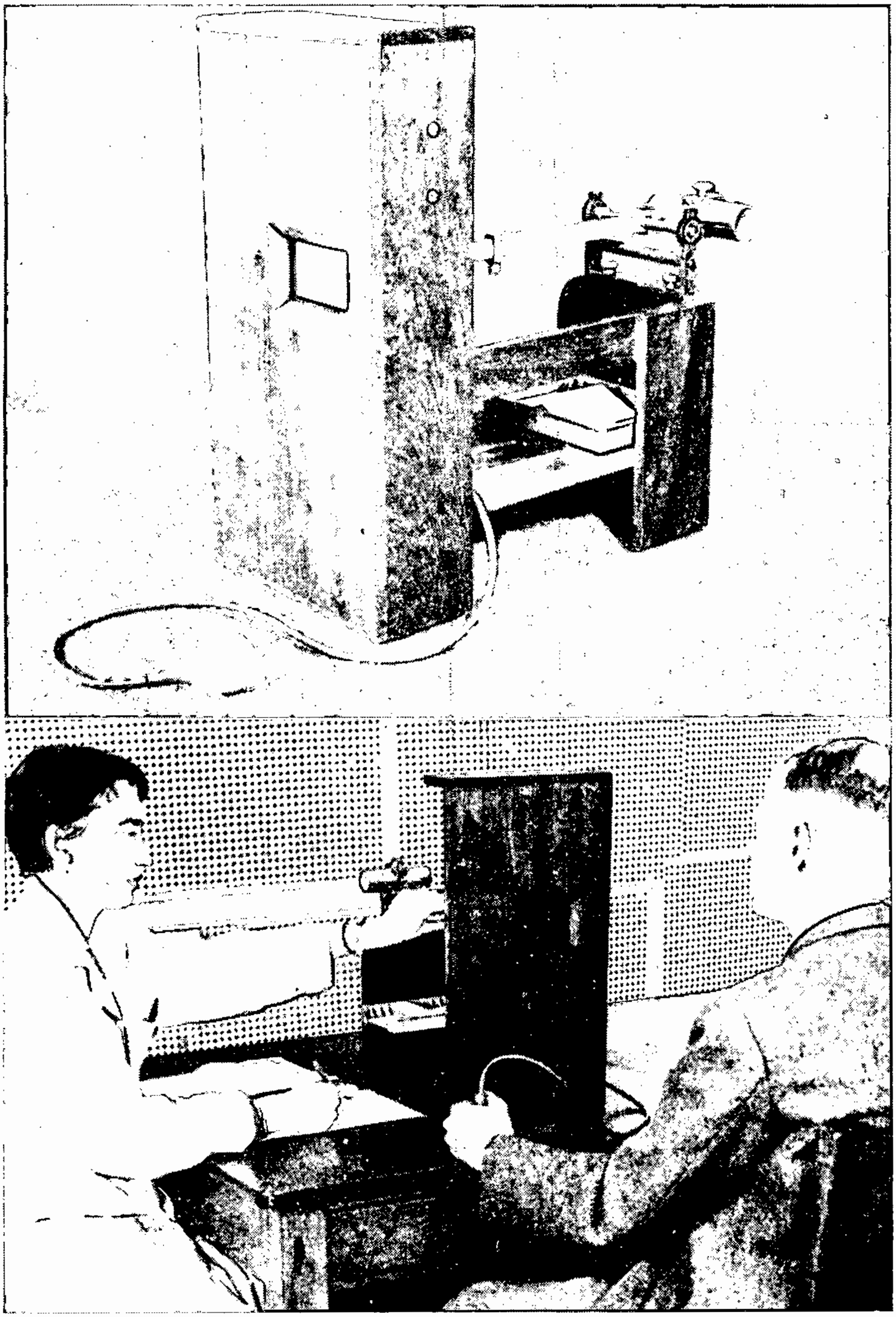


Plate 1

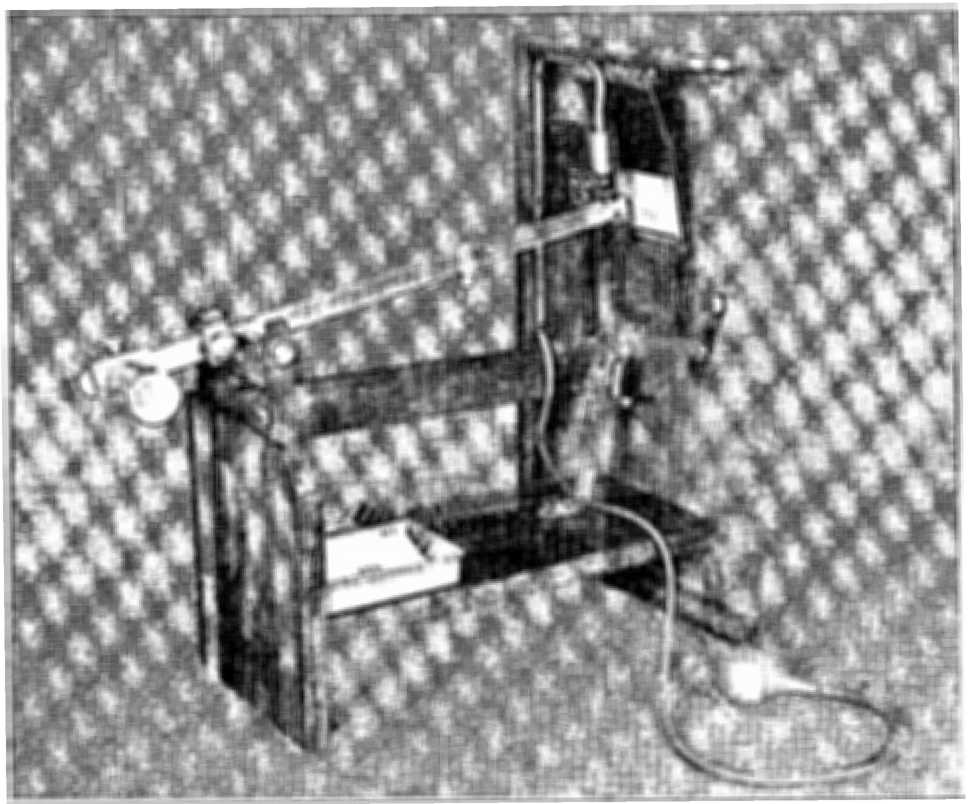


Plate 2

If a " constant-speed " series is being done, the cursor-weight is left at its pre-set position, and the first card simply replaced by the second, which is exposed in the same way.

For the presentation of moving stimuli, procedures are the same, except that the cards are placed in the holder on the segment instead of in the fixed holder.

In adjusting the counterpoise and trimming weights to a suitable zero for a particular series of stimuli, the aim should be an arrangement that allows settings of the cursor-weight along the scale to cover the full range of individual differences which are likely to be encountered in the test sample. The main consideration is the nature of the stimuli, and it is advisable to do some brief initial experimentation on a few heterogeneous subjects.

A possible improvement would be the addition of a holder in a recess near the back edge of the window to take flat metal or card masks with horizontal slots of varying width cut out at the middle and having outside dimensions corresponding to those of the standard card so that the masks could also be inserted in the existing holder on the segment. By using small stimulus figures and exposing them behind slots just wide enough to accommodate them, the upper speed-limit of exposures possible on the apparatus would be raised considerably.

Adaptation for recording the actual speed of exposures could be done by arranging a contact-system between either the beam-arm or the segment and suitably placed stationary points on the chassis, in conjunction with a chronoscope. For certain work it might be convenient to plot a graph of actual times corresponding to positions of the cursor-weight on the linear scale, which would, of course, be specific for the particular range of speeds imposed by the pre-settings of the counterpoise- and trimming-weights and the vertical extent of the stimulus. The plotting of such a graph for this apparatus, using the alternative means of a photo-electric recording system, has been done by Mundy-Castle, McKiever and Alexander (2).

Four reproductions of the dual-purpose " fall " tachistoscope were constructed in the Department of Psychology at the University of the Witwatersrand for use by advanced students. The author has been given to understand that the design has proved reasonably satisfactory, but for certain investigations the unavoidable sound of the catches and the beam-end coming in contact with the stop is a disadvantage.

ACKNOWLEDGMENT

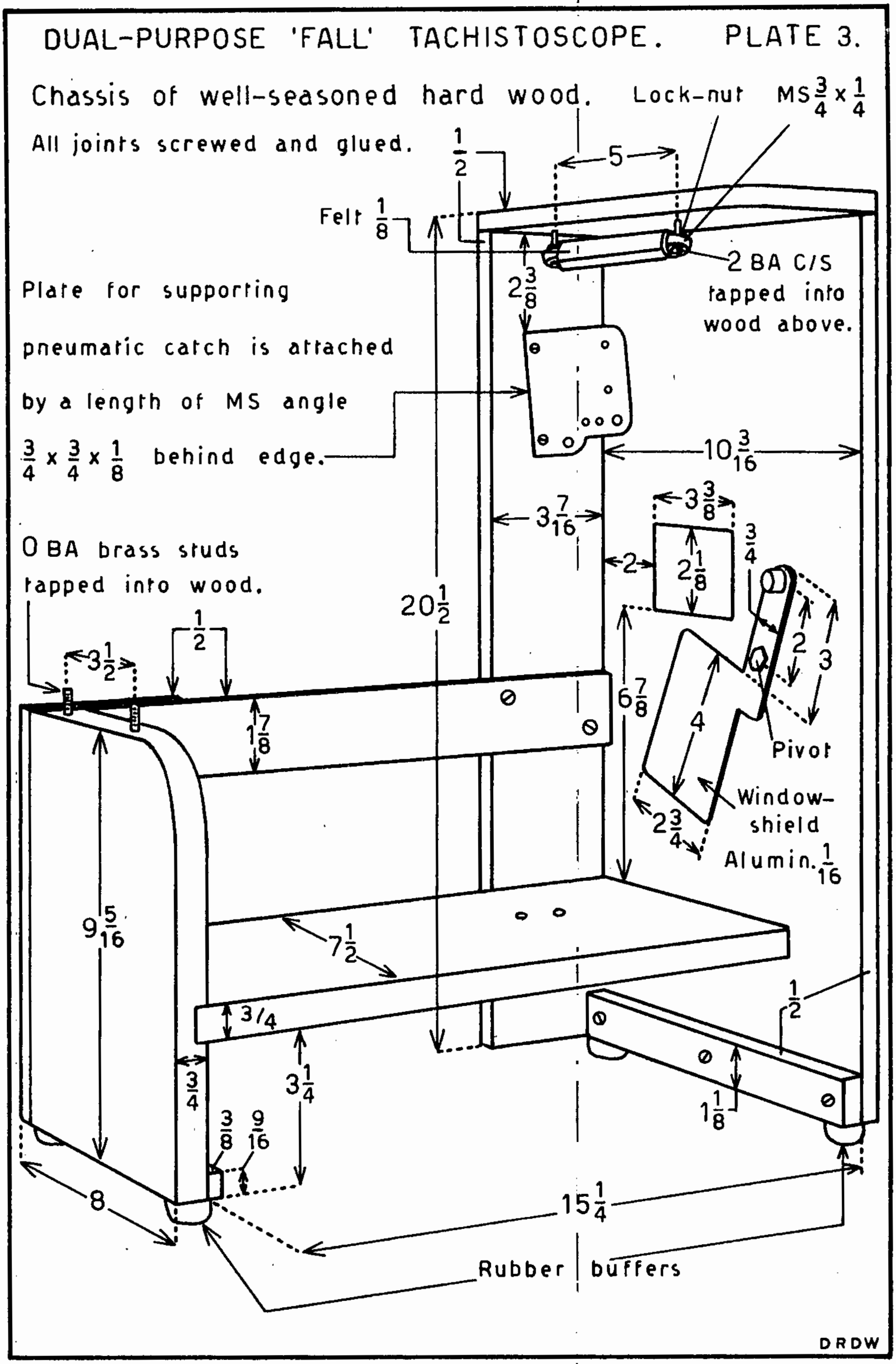
I wish to thank Mr. J. B. Kirstein for his assistance in photographing the apparatus.

REFERENCES

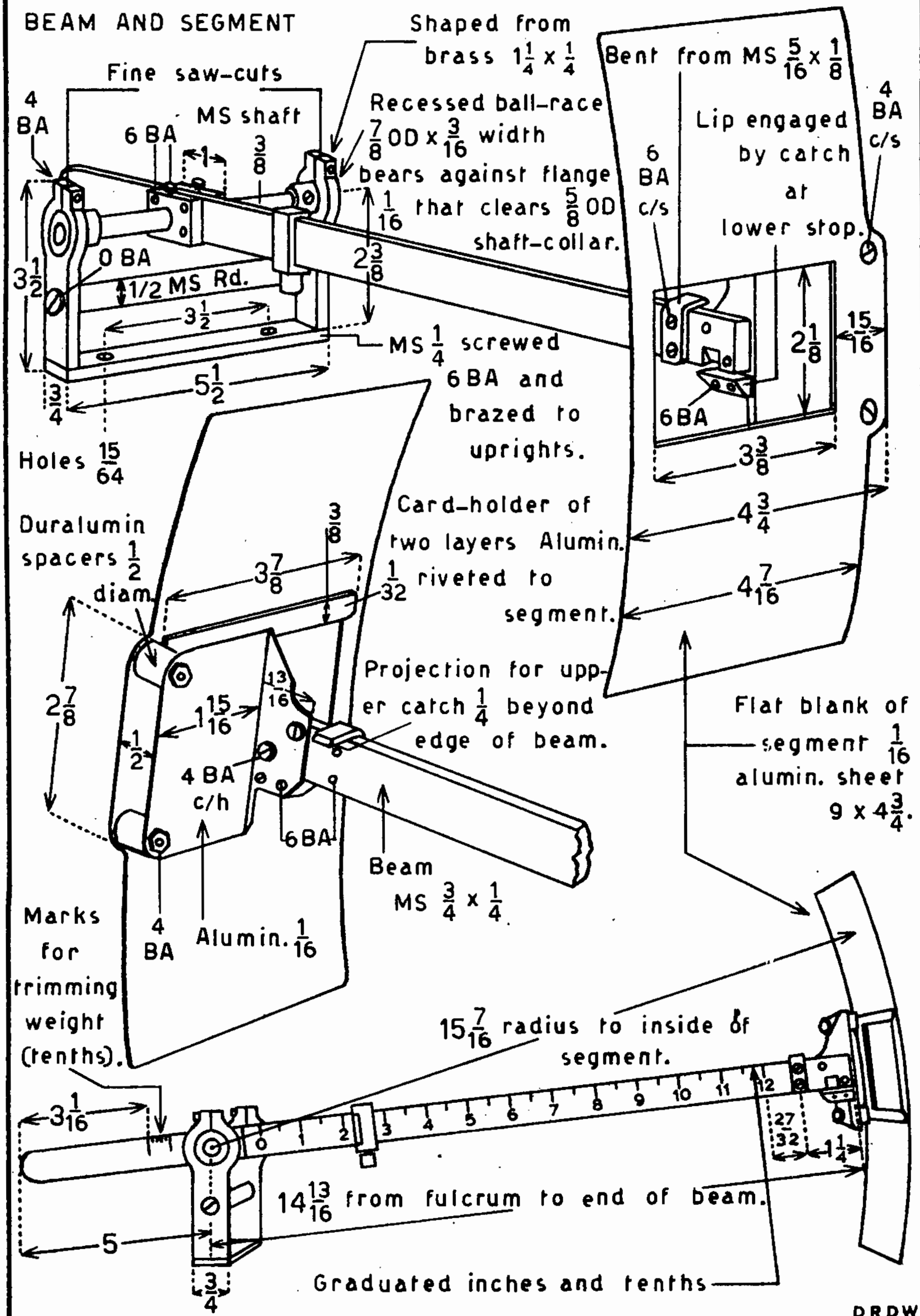
1. BIESHEUVEL, S. *and* A. C. MUNDY-CASTLE. An investigation into the visibility of road signs. *Bull. Nat. Inst. Personnel Res.*, 1952, 4, No. 1, 60–65.
2. MUNDY-CASTLE, A. C., B. McKIEVER *and* V. ALEXANDER. The determination of tachistoscopic exposure time. *Bull. Nat. Inst. Personnel Res.*, 1950, 2, No. 3, 40–43.
3. MUNDY-CASTLE, A. C. The alpha rhythm and rate of visual perception. *J. Nat. Inst. Personnel Res.*, 1955, 6, 38–43.

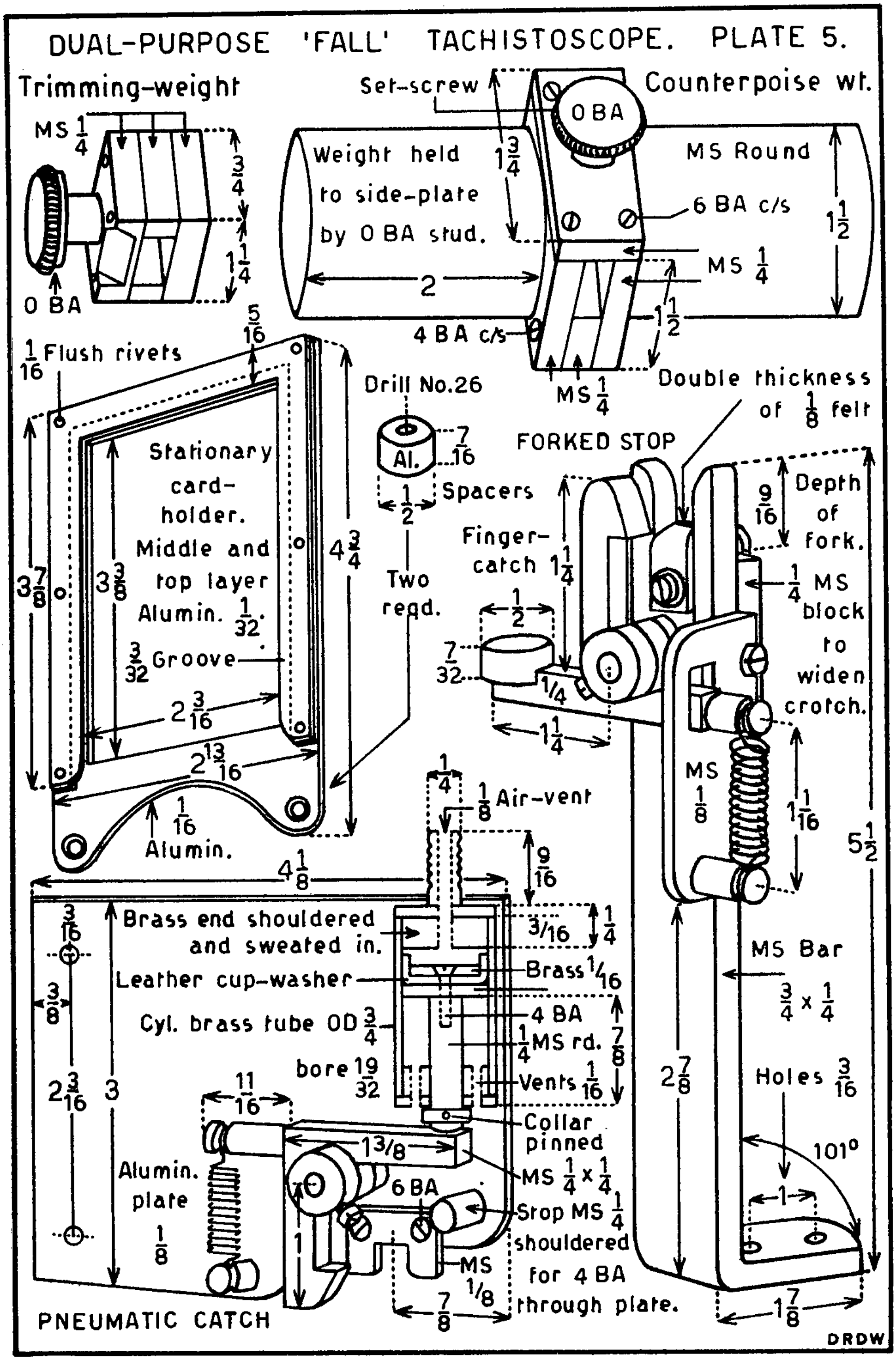
Mr. D. R. de Wet,
National Institute for Personnel Research,
P.O. Box 10319,
Johannesburg,
Republic of South Africa.

Manuscript received: 18th February 1963.
Revised manuscript received: 19th June 1963.


DUAL-PURPOSE 'FALL' TACHISTOSCOPE. PLATE 3.
Chassis of well-seasoned hard wood.
All joints screwed and glued.
Lock-nut
MS 3/4 x 1/4
1/2
5
Felt 1/8
2 BA C/S tapped into wood above.
2 3/8
Plate for supporting pneumatic catch is attached by a length of MS angle 3/4 x 3/4 x 1/8 behind edge.
10 3/16
3 3/8
3 7/16
3/4
2
2 1/8
O BA brass studs tapped into wood.
20 1/2
1/2
3 1/2
6 7/8
2
3
4
1 7/8
Pivot
Window-shield
2 3/4
Alumin. 1/16
9 5/16
7 1/2
1/2
3/4
3 1/4
1 1/8
3/8
9/16
8
15 1/4
Rubber buffers
DRDW

DUAL-PURPOSE 'FALL' TACHISTOSCOPE. PLATE 4.




DUAL-PURPOSE 'FALL' TACHISTOSCOPE. PLATE 5.
Trimming-weight
MS 1/4
3/4
1 1/4
0 BA
Set-screw
0 BA
Counterpoise wt.
Weight held to side-plate by 0 BA stud.
1 3/4
MS Round
6 BA c/s
1 1/2
MS 1/4
2
1 1/2
4 BA c/s
MS 1/4
Double thickness of 1/8 felt
1/16 Flush rivets
5/16
Drill No.26
Al.
7/16
1/2
Spacers
Two reqd.
FORKED STOP
Stationary card-holder. Middle and top layer Alumin. 1/32.
3/32 Groove
3 7/8
3 3/8
4 3/4
2 3/16
2 13/16
1/16 Alumin.
Finger-catch 1 1/4
1/2
7/32
1/4
1 1/4
9/16
Depth of fork.
1/4 MS block to widen crotch.
MS 1/8
1 1/16
5 1/2
1/4
1/8 Air-vent
9/16
4 1/8
3/16
Brass end shouldered and sweated in.
3/16
1/4
Leather cup-washer
Brass 1/16
Cyl. brass tube OD 3/4
4 BA
1/4 MS rd.
7/8
bore 19/32
Vents 1/16
3/8
2 3/16
3
11/16
Collar pinned
1 3/8
MS 1/4 x 1/4
Alumin. plate 1/8
6 BA
Stop MS 1/4 shouldered for 4 BA through plate.
1
MS 1/8
7/8
PNEUMATIC CATCH
MS Bar 3/4 x 1/4
2 7/8
Holes 3/16
101°
1
1 7/8
DRDW

Reprinted from "Psychologia Africana," Volume 10, 1964.

MEASURES OF SPEED OF PERCEPTION AND SPAN OF ATTENTION ON AIR-PILOT CANDIDATES

D. R. DE WET

Using the dual-purpose " fall " tachistoscope (3), two tests, intended respectively, for measuring speed of perception and span of attention, were administered experimentally to a group of air-pilot candidates in 1956 (Table 1), to sound the possible importance of these functions in airmanship and also investigate their correlation with certain other variates—intellectual, sensory-motor, and perceptual. All the candidates were physically fit and had normal eye-sight.

In this particular study, " speed of perception " involved the identification of visual stimuli moving quickly across a restricted field, and " span of attention ", the identification of stationary visual stimuli presented for a short period. (It is appreciated that there are other connotations in which these terms can be used.)

On the basis of their performance on a battery of standardized selection tests, which did not include the above two experimental tests, forty-seven of the total sample (Table 1) were selected to undergo training as air-pilots, and in due course, thirteen of these were suspended. Of the latter group, only the eight candidates suspended for poor flying ability have been included for validation purposes.

Table 1

Derivation of Sample

[Applied for flying course ..	163
[Tested for Speed of Percept. and Span of Att. ..	161
Accepted for training (by other tests) ..	47
Suspended from training (all reasons) ..	13
[Suspended for poor flying ability ..	8
[Tested for Speed of Percept. and Span of Att. ..	8
[Completed course successfully ..	33
[Tested for Speed of Percept. and Span of Att. ..	32

Speed of Perception Test (Procedure)

The material consisted of ten white cards, each one having a stimulus of three symbols printed across the middle in black. Plain Roman capitals and Arabic numerals were used, all being $\frac{9}{32}''$ in height, approximately $\frac{3}{16}''$ in breadth, and $\frac{3}{64}''$ in width of line, each set of three symbols occupying a lateral space of about $\frac{3}{4}''$. The legend on each card and the order of presentation is indicated in Table 3. The counterpoise- and trimming-weights on the tachistoscope were set at suitable positions to cause the stimulus on a card in the segment to fall past the window in about half a second when the cursor-weight was at the first graduation-mark (near the fulcrum); approximate checks being done with a stop-watch. With this basic setting, and the cursor-weight at the tenth mark on the beam, the stimulus took about $^1/_{20}$th of a second to pass across the window.

Having seated the subject at his most comfortable distance from the window for reading ordinary typescript, he was instructed, that when the observer said " Yes ", he was to watch the window carefully, press the rubber bulb, and try to see what was printed on the card that came down.

With the cursor-weight set at the tenth mark, the beam was cocked, and the subject given the sign to make his exposure. If he succeeded in identifying all three symbols correctly, he scored ten points, and proceeded to the second card, which was exposed in the same way, with the cursor-weight at the tenth mark. But if he failed to identify any component correctly, the cursor-weight was moved to the ninth mark and another exposure given. Success this time scored nine points, and failure meant another trial with the cursor-weight at the eighth mark. For each repetition required, the exposure time was increased by setting the cursor-weight an inch nearer the fulcrum, and if the subject still failed to identify the entire stimulus when the cursor-weight was set at mark one, he scored nought. This procedure, of starting with the shortest exposure and gradually increasing it until either the subject's threshold or the lowest mark on the scale was reached, was applied to each of the cards in the series.

Speed of Perception Test (Results)

Test reliability by split-half correlation on the "odd-even" pattern between Cards 1-4-5-8-9 and Cards 2-3-6-7-10 with correction by the Spearman-Brown formula for prediction, is given in Table 2 (first column). It was thought that, in a task involving the identification of stimuli moving across a restricted field, random fluctuations of attention would be very considerable, and markedly reduce the reliability; but in view of the substantial reliability coefficient obtained, this chance influence appears to have been quite small.

Table 2

Product-moment inter-trial correlations

N = 161	Speed of Per.	Span of Att.
Uncorrected ..	0·887	0·683
Corrected ..	0·940	0·812

The mean scores obtained by the whole test group (N = 161) on the respective cards are shown in Table 3. These figures represent average distances in inches of the cursor-weight from the zero-mark near the fulcrum; the larger the figure, the higher the speed-threshold at which the corresponding legend could be perceived, and hence the "easier" the card. Table 4 indicates that the differences between pairs of means are, for the most part, significant at better than the 5 per cent level—there being only three exceptions.

The average for the means is 3·93. Cards 2 (RKE) and 5 (659) are not significantly different from this, whereas Cards 1 (B5R), 9 (S3S), and 10 (CGO) are above the average in difficulty, and Cards 8 (ETL), 7 (7Z7), 6 (HMN), 4 (ZZ7) and 3 (EFP), below the average. There is the suggestion here, that moving configurations which consist principally of curves are more difficult to perceive than those having a pre-

Table 3

Speed of Perception

Mean scores of the whole test sample (N: 161)

Card	Legend	Mean	S.D.
1	B5R	0·93	1·46
2	RKE	3·67	3·06
3	EFP	5·84	3·26
4	ZZ7	4·88	3·37
5	659	3·95	3·39
6	HMN	5·11	3·33
7	7Z7	5·23	3·37
8	ETL	5·94	3·10
9	S3S	1·63	2·59
10	CGO	2·10	2·56
Average		3·93	

Table 4

Speed of Perception

Significance of the differences between the means

Card	1	2	3	4	5	6	7	8	9	10
1										
2	S									
3	S	S								
4	S	S	S							
5	S	S	S	S						
6	S	S	S	—	S					
7	S	S	S	S	S	—				
8	S	S	—	S	S	S	S			
9	S	S	S	S	S	S	S	S		
10	S	S	S	S	S	S	S	S	S	

" S " implies that difference is significant at better than the 5 per cent level.

dominence of straight lines. In the case of the latter, blurring through movement seems to have less effect in obscuring the basic cues for identification. On this hypothesis, Card 1 (B5R) could be expected to be easier to identify than either Card 9 (S3S) or Card 10 (CGO), but actually it shows up as rather more difficult, because, being the first card in the series, it imposed the additional task on the subjects of adapting themselves to an unfamiliar test situation. (Considerable research has been done in connection with the comparative legibility of *stationary* letters by Hartridge and Owen (6), Woodruff (14), and others. In general, the findings suggest that letters in which curves or diagonal straight lines predominate are less legible than " rectangular " letters.).

On the present data it is not possible to demonstrate any systematic effect of practice, but it probably does play a substantial part, particularly in the earlier trials. Card 4 (ZZ7) might well be inherently more difficult than Card 7 (7Z7), but on the other hand the legends are structurally very similar, and it seems more likely that subjects found Card 7 easier because of increased practice at this stage in the series.

Table 5 gives the respective mean total scores of the pupils who succeeded on the course of flying training and those who were suspended from training because of poor airmanship. The lack of any significant difference indicates that speed of perception as measured in the present test is unrelated to success or failure in learning to fly.

Table 5

Speed of Perception and Flying Training

Pupils	N	Mean Total	S.D.
Successful	32	43·72	23·28
Suspended	8	44·25	20·34
(No significant difference between means)			

Span of Attention Test (Procedure)

The ten cards used in this test were inscribed with black capital letters and numerals similar in size and format to those in the other; the first having three components; the second and third, four; the fourth to the seventh, inclusive, five; the eighth and ninth, six; and the tenth, seven. After adjusting the weights on the tachistoscope to give constant exposures of about $^1/_{25}$th of a second, and arranging the subject at his most comfortable " reading " distance from the window, each card was inserted in turn in the stationary holder and exposed when the subject pressed the bulb. As before, he was told to watch carefully and report what he saw, only one trial being given with each card and a point scored for every symbol correctly identified.

Span of Attention Test (Results)

As shown in Table 2 (second column), by the split-half correlation on the " odd-even " pattern between the two groups of trials, corrected by the Spearman-Brown formula, the reliability of this test is somewhat lower than that of " speed of perception ", but still quite acceptable. A likely explanation for the lower reliability, lies in the fact that only a single trial on each card was given, whereas in the previous test, trials were repeated, if necessary, until the subject perceived the stimulus or failed at the tenth trial. Thus, contrary to what was expected, random fluctuations of attention probably played a bigger part in the Span of Attention test—because of the particular technique of measurement employed.

Table 6 gives the frequency distributions of the number of symbols missed on each card. Those for the last three cards, which have a comparatively large number of symbols, are less markedly skewed towards the " easy " side than the rest. A symbol was regarded as " missed " if the subject failed to report it or reported it incorrectly.

Table 6

Span of Attention

Frequency distributions of the number of symbols missed

Card number and Legend		Number of symbols missed							
		0	1	2	3	4	5	6	7
1 **G59**	N %	62 38	88 55	8 5	3 2				
2 **1S38**	N %	53 33	97 60	9 6	2 1	0 0			
3 **T358**	N %	142 88	18 11	1 <1	0 0	0 0			
4 **2ZV43**	N %	31 19	52 32	54 34	22 14	2 1	0 0		
5 **6DB9V**	N %	9 6	57 35	61 38	31 19	2 1	1 <1		
6 **H72X5**	N %	81 50	48 30	24 15	7 4	0 0	1 <1		
7 **R56G7**	N %	18 11	111 69	29 18	2 1	1 <1	0 0		
8 **1KTZ75**	N %	1 <1	31 19	79 49	37 23	12 7	1 <1	0 0	
9 **N6RP31**	N %	2 1	26 16	67 42	53 33	13 8	0 0	0 0	
10 **3TG5U42**	N %	3 2	21 13	37 23	54 33	39 24	6 4	1 <1	0 0

The total frequency with which each of the symbols constituting the stimulus on each card was missed is shown in Table 7. On Card 1 (G59), the commonest error was to mistake the " G " for the numeral " 6 ", and on Card 2 (1S38), the " S " was frequently reported as an " 8 ", and the " 3 " very occasionally as an " 8 ", but in no instance was the numeral " 1 " mistaken for the capital letter " I ". On Card 3 (T358) there is a marked reduction in the proportion of missed symbols, probably as a result of practice and increased familiarity with the test situation. Those subjects who missed the " T ", reported it as the numeral " 1 ". On the other cards, the most noteworthy cases of mistaken identity were " Z " reported as " 7 "; " V " as " Z ", " 7 ", or " X " (the multiplication sign more often than the letter); " B " as " 8 "; " G " as " 6 "; " T " as " 1 "; " U " as " IJ " or double " I ". (On Card 6, subjects scored a point whether they reported the " X " as a letter or the multiplication sign " times ".) Symbols missed in the larger stimuli of five items or more were usually not reported on, but throughout the series, it was observed that when symbols were wrongly reported it was more common for certain alphabetical letters to be mistaken for numerals than numerals for letters.

Table 7

Span of Attention

Total frequency with which each symbol was missed

Card	Symbols and total missed						
1	G 91	5 9	9 13				
2	1 0	S 100	3 14	8 7			
3	T 9	3 3	5 3	8 5			
4	2 10	Z 26	V 73	4 60	3 65		
5	6 3	D 14	B 117	9 98	V 53		
6	H 2	7 12	2 31	X 46	5 31		
7	R 3	5 6	6 14	G 133	7 23		
8	1 7	K 5	T 45	Z 146	7 114	5 36	
9	N 2	6 11	R 40	P 108	3 129	1 81	
10	3 8	T 7	G 38	5 88	U 101	4 108	2 99

Table 8 gives the respective percentage frequency of letters and numerals missed, i.e. wrongly reported or not reported at all. Notwithstanding the fact that there is generally a greater loading of numerals than letters at the right-hand ends of the stimuli, which, because of the usual habit of " left-to-right " orientation for this kind of material could be expected to weight the proportion of missed symbols in favour of the numerals (as it actually does, very markedly, in the case of the last two cards), there is still a significant overall tendency for a larger proportion of letters than numerals to be missed ($P < 0 \cdot 01$).

Some interesting questions arise that cannot be answered on the strength of the present data: Is this apparent difference in perceptual threshold, coupled with the tendency for certain letters to be mistaken for numerals, rather than the reverse, an indication that, in general, numerals are more fundamental than letters, or is it a local occurrence in a test-sample " selected " in the sense that persons aspiring to become air-pilots are mainly those whose interests favour the mathematical and scientific and who might therefore have some particular bias towards numerical concepts? Are numerals more " familiar " than letters because there are only ten distinct configurations amongst the numerals as compared with twenty-six in the

Table 8

Span of Attention

Respective percentages of letters and numerals missed

Card	Letters (%)	Numerals (%)
1	56·52	6·83
2	62·11	4·35
3	5·59	2·28
4	30·75	27·95
5	38·10	31·37
6	14·91	15·32
7	42·24	8·90
8	40·58	32·51
9	31·06	45·76
10	30·23	47·05
Mean	35·21	22·23
P of the Difference $<0{\cdot}01$		

alphabet and numerals are usually learned before letters? Is this familiarity reinforced by the fact that each numerical symbol by itself represents a meaningful unit, whereas the single letter does so comparatively seldom—usually it takes a group of two or more to produce a conceptual " gestalt " in the form of a word?

In an early experiment, Jacobs (7) found that school-children could repeat, after utterance by the observer, an average of 6·1 nonsense syllables, 7·3 letters, and 9·3 numerals. He points out that nonsense syllables, being " new to the hearer, have to be learnt, and absorb more energy ", and that " our expectant attention " has to search among fewer numerals than letters.

With regard to the perception of words, Postman, Bruner and McGinnies (8) have shown that words relating to matter which interests the subject are more easily apprehended than those unconnected with his interests. Solomon and Howes (11) hold the view that familiarity, whether it derives from personal interest or not, is a factor that makes certain word-stimuli more easily recognizable. It seems reasonable to suppose that the influence will be stronger when both interest and familiarity are present, and that this applies, not only to the " meaningfulness " of word-stimuli, but also to the actual symbolic medium, such as letters or numerals.

Table 9 reflects the percentage of each stimulus missed by the test-group as a whole, and the significance of the differences between adjacent cards. In Table 10, the signs " S " or " - " indicate respectively, whether the difference between each card and every other card is significant at better than the 5 per cent level, or not. Referring to Table 9, the most easily perceived stimulus is that on Card 3 (T358). As it appears to be structurally no simpler than the two preceding cards, a likely reason is practice effect, which, as Weber (13), Renshaw (9), and others have shown, can produce a marked improvement in the perceptual span. For the same reason, Card 2 (1S38), which could be expected to be more difficult than Card 1 (G59), because it has an additional symbol, is not significantly more difficult, and Cards 6 and 7 tend to be easier than Cards 4 and 5, although they all consist of five symbols. Furthermore, Card 10 (seven symbols) is not significantly more difficult than Cards 8 and 9 (six symbols), or Card 5 (five symbols).

Table 9

Span of Attention

Percentage of the entire stimulus missed on each card

Card	Legend	Per cent	t of Diff.
1	G59	23·40	1·39
2	1S38	18·79	5·81
3	T358	3·11	6·84
4	2ZV43	29·07	1·66
5	6DB9V	35·40	5·61
6	H72X5	15·16	2·19
7	R56G7	22·24	3·97
8	1KTZ75	36·54	0·50
9	N6RP31	38·41	0·36
10	3TG5U42	39·84	

(t > 1·96 implies significance at better than the 5 per cent level)

Table 10

Span of Attention

Significance of the differences

between the percentages of all symbols missed

Card	1	2	3	4	5	6	7	8	9	10
1										
2	—									
3	S	S								
4	—	S	S							
5	S	S	S	—						
6	S	—	S	S	S					
7	—	—	S	—	S	S				
8	S	S	S	S	—	S	S			
9	S	S	S	S	—	S	S	—		
10	S	S	S	S	—	S	S	—	—	

" S " implies that difference is significant at better than the 5 per cent level.

Referring again to the total frequencies of the missed symbols on Cards 2, 5, 6, 7, 8, 9, and 10 (Table 7), it is noteworthy that on each of these cards the frequency of misses for the last symbol, while greater than that for the first, is less than that for the second last; and on Cards 2, 4, 5, 8, 9, and 10, the frequency for the last symbol is less than that for the third last (although, on Card 10 the difference is hardly significant).

It appears, therefore, as though the left and right " boundary " components of the stimulus have some special strength relating to their physical position. The weakest sections (where the most misses occur) incline towards the right, through the normal habit of starting to read material of this sort from left to right, but they do not occur at the extreme right. This suggestion of a " closure " effect, whereby stimuli compounded of discrete units are perceived to a considerable extent as wholes, even when these wholes, as such, are not familiar or meaningful, is in line with the findings of Wagner (12) and Schumann (10) who exposed rows of letters and instructed their subjects to attend to the whole series each time.

Although, in the present test, the fixation point of attention had a sinistral bias, as the subjects were free to follow the usual habit of left-to-right orientation, there is still evidence of some propensity to sacrifice inner detail in favour of grasping more of the overall configuration. This becomes less apparent when the stimulus consists of as many as seven unconnected but individually familiar items (Card 10), which Glanville and Dallenbach (5) found was about the average span for such material. In the case of a stimulus constituting a *meaningful* unit, the closure-effect would be very much greater, as originally demonstrated by Cattell (1) and confirmed by Erdmann and Dodge (4) who found that familiar words of up to twenty letters could be correctly identified at a very brief exposure.

The relationship between the Span of Attention test and achievement on the flying course is shown in Tables 11 and 12. There is a significant tendency for those pupils who completed the course successfully to have a better span than those who were suspended for poor flying ability. This is indicated by a significant difference between the means of the two groups, and a biserial correlation of $0 \cdot 48$ ($P < 0 \cdot 01$) between the test and the Success/Suspension criterion. Learning to pilot an aircraft, which is a fast and complex machine fitted with numerous controls and instruments, often requires the apprehension of several things almost simultaneously.

Table 11

Span of Attention

Distributions of scores on the pupil group

Score	Total	Successes	Suspensions
16	1	0	1
28	2	2	0
29	1	1	0
32	2	1	1
33	2	1	1
34	3	1	2
36	3	2	1
37	5	4	1
38	7	7	0
39	4	4	0
41	4	4	0
42	1	1	0
44	2	1	1
45	2	2	0
47	1	1	0
N	40	32	8

Table 12

Span of Attention and Flying Training

Pupils	N	Mean Total	S.D.	P of Diff.
Successful	32	37·91	4·42	<0·03
Suspended	8	33·25	7·40	
Biserial r between test and Success/Suspension: 0·48 (P<0·01)				

The distributions of the total scores obtained on the Speed of Perception and Span of Attention tests are given in Table 13. The former is platykurtic, and has a bias towards " difficulty "; the latter, more normal, but skewed towards the " easy " side.

Table 13

Distributions of Total Scores

Speed of Perception		Span of Attention	
Score	N	Score	N
0— 6	11		
7—13	11	16—18	1
14—20	16	19—21	0
21—27	17	22—24	1
28—34	21	25—27	5
35—41	16	28—30	14
42—48	15	31—33	21
49—55	12	34—36	34
56—62	11	37—39	56
63—69	7	40—42	19
70—76	11	43—45	9
77—83	6	46—48	1
84—90	7		
	161		161

In deriving the correlations between these two tests and the various others, as shown in Table 14, raw scores were used throughout. Both correlate to a small but significant extent with Raven's Progressive Matrices (Non-verbal intelligence), and slightly more with the learning section on the Hand-foot Reaction test (2), which, in addition to its sensory-motor element, involves a certain amount of non-verbal intelligence. (The correlation between the R.T. learning section and Raven's test, not shown in the table, is 0·278.) Span of Attention is also related to Mental Alertness (Verbal intelligence), Arithmetic, Technical and Scientific Information, and the more specific knowledge of matters pertaining to aircraft and flying; and Speed of Perception, to speed on the Hand-foot Reaction test, and less error on the Air-pilots' Arm-leg co-ordination test and the Two-hand co-ordination test (Moede type). In general, Speed of Perception is associated more with certain of the sensory-motor activities, and Span of Attention with certain of the intellectual. The relationships, however, are small. There is no significant correlation between Speed of Perception and Span of Attention.

Table 14

Product-moment correlations between Speed of Perception, Span of Attention, and other Tests (N: 161)

(Decimal points have been omitted)

Other Tests	Speed of Per.	Span of Att.
Mental Alertness	127	251 S
Progressive Matrices (Raven's)	163 S	196 S
Arithmetic	012	196 S
Mathematical Achievement	121	123
Technical and Scientific Information	138	184 S
Tech. and Sc. Reading Comprehension	129	127
Flying Interest Inventory	069	185 S
Air-pilots' Arm-leg Co-ord. (Error)	−221 S	−116
Air-pilots' A.L.C.-Distraction (Error)	−180 S	−104
Manual Steadiness (Total Time)	024	−132
Manual Steadiness (Total Error)	003	097
Hand-foot Reaction (Learning Time)	−291 S	−220 S
Hand-foot Reaction (Speed of response)	163 S	084
Hand-foot Reaction (Total Error)	107	087
Two-hand Co-ord., Moede Type (Error)	−247 S	−112
Two-hand Co-ord., Moede Type (Time)	−137	−113
Foveal Flicker Fusion	−112	−049
Product-moment correlation between Speed of Perception and Span of Attention: 125		
" S " implies significance at better than the 5 per cent level		

Summary and conclusions

1. By means of a dual-purpose " fall " tachistoscope (3), 161 air-pilot candidates were tested for Speed of Perception, involving the identification of visual stimuli moving quickly across a restricted field, and Span of Attention, involving the identification of stationary visual stimuli exposed for a brief period.

2. The ability to identify either moving or stationary stimuli is improved by practice.

3. Contrary to expectation, the Speed of Perception test showed somewhat better inter-trial reliability than the Span of Attention test, which, nevertheless, showed adequate reliability.

4. Moving stimuli that consist mainly of curves seem to be more difficult to perceive than those made up principally of straight lines.

5. Speed of Perception, as measured by the present test, does not relate to success or failure in learning to fly.

6. A biserial correlation of 0·48 ($P < 0\cdot01$) obtained between performance on the Span of Attention test and Success/Suspension in flying training, implies that good Span of Attention is associated with success in flying.

7. Frequency distributions obtained with stimuli of less than six components tend to be skewed towards the easy side.

8. It was observed that when symbols were wrongly reported, it was more common for certain letters to be incorrectly identified as numerals, than the reverse.

9. There is also an overall tendency for a larger proportion of letters than numerals to be " missed ", i.e. not reported at all, or wrongly reported. ($P < 0{\cdot}01$.)

10. In line with the results of other investigators, there is evidence that stimuli compounded of several discrete units are perceived to a considerable extent as wholes, in spite of the fact that these wholes, as such, are neither familiar nor meaningful.

11. Correlations between the Speed of Perception and Span of Attention tests and other variates are small. They do, however, suggest that Speed of Perception relates more to physical reaction and co-ordination, and Span of Attention to intellectual activity.

12. There is no significant correlation between Speed of Perception and Span of Attention, as measured by the present tests.

ACKNOWLEDGMENTS

I am indebted to Mr. M. A. Coulter, Mr. D. Humphriss and Dr. H. Reuning who reviewed the manuscript and suggested several improvements.

REFERENCES

1. CATTELL, J. McK. Ueber die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben. *Phil. Stud.*, 1885, 2, 635–650.
2. DE WET, D. R. A portable hand-foot reaction test. *J. Nat. Inst. Personnel Res.*, 1960, 8, 106–116.
3. DE WET, D. R. A dual-purpose " fall " tachistoscope. *Psychol. Afr.*, 1963, 10, 151–158.
4. ERDMANN, B. *and* R. DODGE. *Psychologische untersuchungen über das Lesen.* Halle, M. Niemeyer, 1898.
5. GLANVILLE, A. D. *and* K. M. DALLENBACH. The range of attention. *Amer. J. Psychol.*, 1920, 41, 207–236.
6. HARTRIDGE, H. *and* H. B. OWEN. Test types. *Brit. J. Ophthal.*, 1922, 6, 543–549.
7. JACOBS, J. Experiments on " Prehension ". *Mind*, 1887, 12, No. 45, 75–79.
8. POSTMAN, L., J. S. BRUNER *and* E. McGINNIES. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, 43, 142–154.
9. RENSHAW, S. The visual perception and reproduction of forms by tachistoscopic methods. *J. Psychol.*, 1945, 20, 217–232.
10. SCHUMANN, F. Das Erkennungsurteil. *Z. Psychol.*, 1921, 88, 205–224.
11. SOLOMON, R. L. *and* D. H. HOWES. Word frequency, personal values, and visual duration thresholds. *Psychol. Rev.*, 1951, 58, 256–270.
12. WAGNER, J. Experimentelle Beiträge zur Psychologie des Lesens. *Z. Psychol.*, 1918, 80, 1–75.
13. WEBER, C. O. Effects of practice on the perceptual span for letters. *J. gen. Psychol.*, 1942, 26, 347–351.

14. WOODRUFF, E. W. Visual acuity and the selection of test letters. (Paper presented at the London Refraction Hospital Jubilee Congress, Oct. 1947.)

Mr. D. R. de Wet,
National Institute for Personnel Research,
P.O. Box 10319,
Johannesburg,
Republic of South Africa

Manuscript received: 31st March 1964.
Revised manuscript received: 21st May 1964.

PRINTED BY CAPE & TRANSVAAL PRINTERS LIMITED, JOHANNESBURG

SECTION IV

Summary

1. Constructional and research data are given, relating to three visual-perception devices, namely, a specific Locus-estimation Test for selecting ground-control interceptor operators; a Cancellation Timer for administering various printed Cancellation tests; and a Dual-purpose "Fall" Tachistoscope for presenting various visual tasks requiring the exposure of either moving or stationary stimuli for short periods.

2. Performance in Locus-estimation correlated 0.40 (significant at the 7% level) with the operational criterion. As the most specific test among others that were all of the pencil-and-paper or pictorial kind, it made a unique contribution to battery validity. No data on its internal consistence or Test-re-test reliability is presently available. The functioning principle of this apparatus might be adaptable in other equipment where an easily controlled apparent movement is desired.

3. Using the Cancellation Timer, a Dots-cancellation test, primarily perceptual, and measuring "sustained attention" was administered to air-pilot candidates. As in the case of more decidedly "motor" tasks, the "Time" measure had greater reliability than the "Error" measure, which, however, also had acceptable reliability.

4. Neither speed nor accuracy in Dots-cancellation was related to flying achievement, probably because the sustained attention measurement by this test is rather specifically "clerical".

5. All subjects improved significantly in their speed of work, but only those who were initially very inaccurate, improved in accuracy. Errors of omission to cancel occurred nearly twice as often as the more "serious" errors of wrong cancellation. To a barely significant extent the fast workers tended to be less accurate. Initially poor performers showed the most marked improvement in relation to their own starting level.

6. Speed of cancellation correlated moderately with speed in a test of Serial discrimination, speed in two tests of Co-ordination, and also error-proneness in one of the latter. Slightly more substantial correlations were obtained on the Cancellation error measure with error in Serial discrimination and achievement in two intellectual tests.

7. The main significance of this Cancellation study is the additional evidence it provides of parallelism between the performance curves of subjects who are initially good, median and poor; not only in tasks that are predominantly "motor", but also in those where the perceptual element is the most important.

8. Tests of Speed of Perception, involving the identification of visual stimuli moving quickly across a restricted field, and Span of Attention, involving the identification of

stationary visual stimuli exposed for a brief period, were administered to air-pilot candidates on the Dual-purpose "Fall" Tachistoscope. Both tests showed adequate inter-trial reliability; Speed of Perception, with more trials, being the more reliable. Speed of Perception, as measured here, did not relate to achievement in learning to fly, but Span of Attention correlated 0.48 ($p < 0.01$) with Success/Suspension in flying training, which implies that good Span of Attention is associated with success in flying.

In Span of Attention, frequency distributions obtained with stimuli of less than six components were skewed towards the easy side. Stimuli made up of several discrete units tended to be perceived to a considerable extent as wholes, even when the latter, as such, were neither familiar nor meaningful.

Correlations between the Speed of Perception and Span of Attention tests and other variates, while small, suggest that Speed of Perception relates more to physical reaction and co-ordination, and Span of Attention to intellectual activity. There was no significant correlation between the Speed of Perception and Span of Attention tests. In both, the ability to identify the stimuli was improved by practice.

SECTION V

AN APPARATUS OF OBJECTIVE TEMPERAMENTAL SIGNIFICANCE

Reprinted from "Psychologia Africana," Volume 9, 1962.

Psychologia Africana, 1962, 9, 100–118

A COMPACT FLICKER-FUSION MACHINE AND ITS APPLICATION TO AIR-PILOT CANDIDATES

D. R. de Wet

Received 16th May, 1960

Until the beginning of the present century, interest in flicker-fusion centred almost entirely around its importance as a function of visual perception. There was unanimous recognition that at low frequency a regularly intermittent light appeared as separate flashes, and as the frequency increased the impression changed to coarse flicker, then fine flicker, and lastly, a stationary light. Any increase of frequency beyond this point (the CFF or critical fusion-frequency) produced no further change whatever.

The explanation was a lag or inertia in the retina that limited its rate of resolution by necessitating a certain minimum time both for the incoming datum to make an impression and for this impression to be completely effaced before a fresh one registered. As the retinal receptors then, like all other organic material, cannot react instantaneously, it follows that beyond the critical speed there must inevitably be an overlap between the new stimulus and the one already in registration.

While individual differences in this fusion-threshold had been observed, it was not until the famous experiment by Wiersma (12) that there appears to have been any considerable notion that it might also have wider psychological significance outside the field of visual perception.

Wiersma (12) was one of the first to appreciate the need for experimental substantiation of the extremely plausible theory of "inertia" or "secondary-function" advanced by Gross (4) and others to explain certain personality differences. In Wiersma's experiment, three groups of subjects, viz., Manic, Normal and Melancholic, were all given the following three tests: (1) the critical speed at which fusion to grey was perceived on a revolving colour-wheel with two suitably matched sectors, one red, and the other green; (2) the time taken in light-adaptation; (3) the time taken for the recovery of tactile sensitivity after over-stimulation by a strong electric current.

In all tests there were very significant and consistent differences between the means for the three groups. The Manic group required faster rotation of the colour-wheel than the Normal before they obtained fusion, and the Melancholic slower than the Normal. In the two adaptation tests, the Manic group showed the fastest recovery rate, and the Melancholic the slowest.

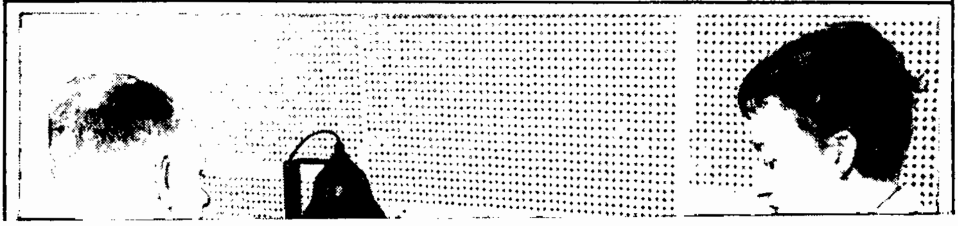
This investigation is particularly noteworthy as a forerunner supplying empirical evidence that individual differences in the sense-modalities, apart from their specific interest, might bear an important relationship to differences in certain aspects of mental integration as a whole.

Much basic work has since been done on the visual CFF (Landis, 5) and definite information garnered about the physical variables such as subject's degree of light-adaptation, his age, the portion of his retina stimulated, his distance from the stimulus, the area and composition of the latter, etc.—all of which influence the limen. However, one of the main attractions of the phenomenon

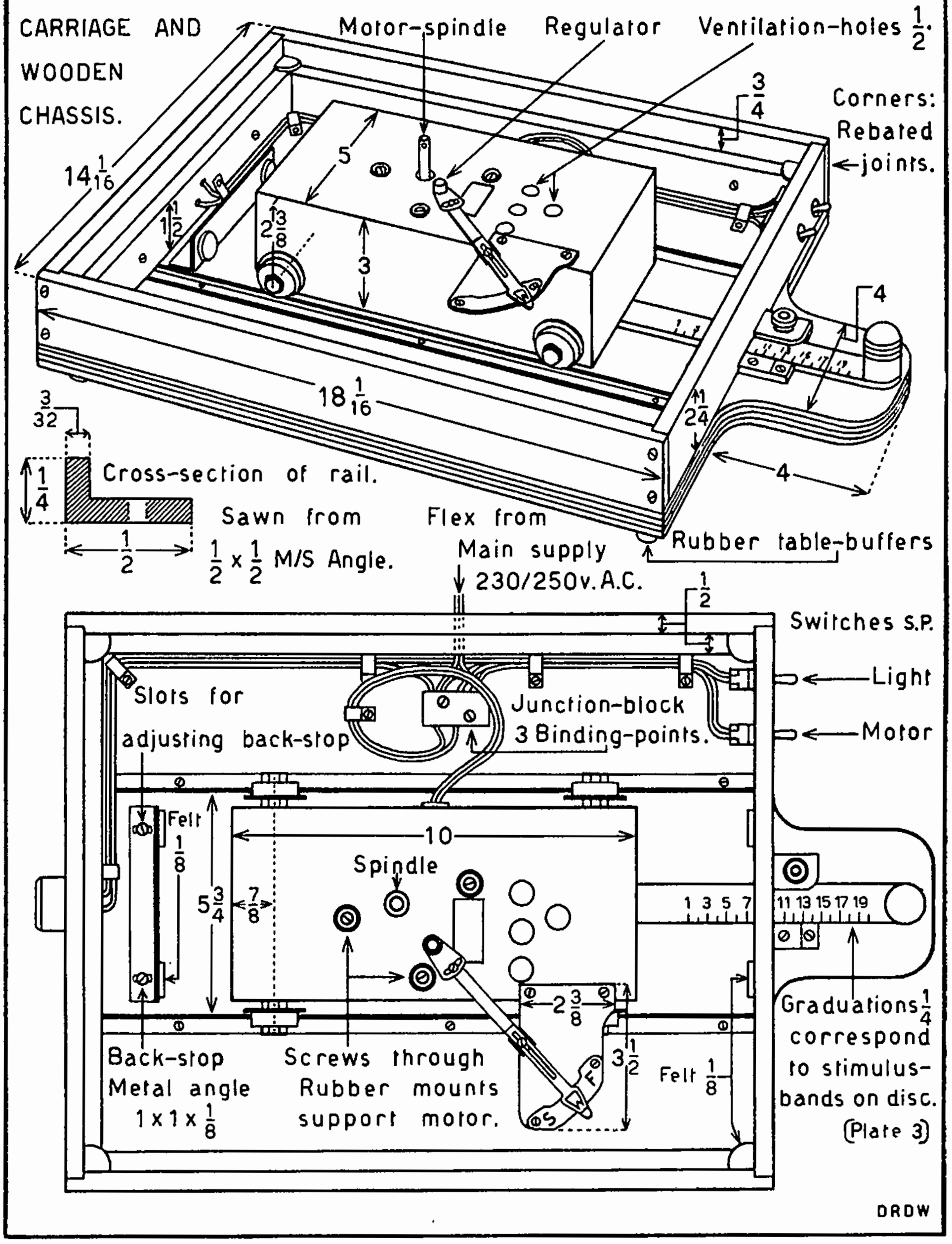
Most apparatus for measuring the visual CFF falls into two broad classes: Electro-mechanical and Electronic. The latter has hitherto proved of somewhat restricted value as it possesses two unsatisfactory features: changes in frequency are usually accompanied by characteristic changes in the sound of the apparatus, and the quality of illumination varies with different frequencies. These drawbacks were certainly experienced with a standard industrial stroboscope in the Aptitude Test Section (A.T.S.) of the South African Air Force during 1943, and have also been noted elsewhere on similar devices.

Electro-mechanical apparatus is of two types, depending on whether motivation is provided by a variable- or constant-speed motor. The following list briefly illustrates some systems.

COMPACT FLICKER-FUSION MACHINE. PLATE 1.



COMPACT FLICKER-FUSION MACHINE. PLATE 2.



1. Variable-speed motor

(a) *Calibrated rheostat*

The light-beam is interrupted by a rotary shutter coupled to the motor-spindle, and variations in frequency are obtained by settings on a calibrated rheostat controlling the motor-speed. This method is identical to that employed in the time-honoured laboratory " colour-mixer," and is subject to the same weaknesses, viz., fluctuations in the supply voltage and also changes in the resistance as the latter warms up, make the correspondence between pre-determined settings on the rheostat and actual motor-speeds, very elastic and unreliable. Beyond its simplicity there is little to recommend this method.

(b) *Coupled speedometer*

Here the arrangement resembles (a) in so far as motor-speed is varied by a controlling rheostat, but there is the important difference that settings are no longer read off on the rheostat. Instead, an additional device coupled to the motor-spindle indicates its speeds directly as number of revolutions per unit time. This is perhaps the most well-known system, and far superior to (a) because, within its limits of accuracy, the speedometer indicates actual motor-speed at all times, irrespective of any uncontrolled fluctuations in voltage and resistance. Disadvantages (two of which are also present in " a ") comprise tell-tale noise accompanying changes in speed, lag between rheostat-control and the response of the motor that hinders quick, nice setting, and the inadequate sensitivity and accuracy of most available speedometer devices.

On an apparatus designed during the last war by Prof. Bozzoli (University of the Witwatersrand) for Dr. Biesheuvel in the A.T.S., speed of rotation was cleverly measured in terms of the E.M.F. produced by a small A.C. bicycle dynamo coupled to the motor-spindle. A suitable series-tuned circuit enabled the voltage generated to be shown on a meter with remarkable accuracy. Unfortunately this plan introduced an additional stage of lag, namely, between motor and recording-meter. On another model, P. Webster of the A.T.S. effected some improvement by substituting a chronometric engine-speed indicator of aircraft pattern that gave more direct measures. This was retained on a later version by the author, but conventional rheostats were replaced by a specially constructed one having a long manual lever that made both fine and coarse adjustments easier and quicker. Although this modification in no way altered the basic inertia of the system, it did afford the observer somewhat better control.

The most promising solution to the main objections outlined above. seems to lie in the use of the constant-speed motor. Some basic experimentation was done by the author in the A.T.S. and later in this Institute on the two approaches " a " and " b " that follow, and although neither materialized into an applied test, both confirmed the existence of good possibilities in the " constant-speed " method.

2. Constant-speed motor

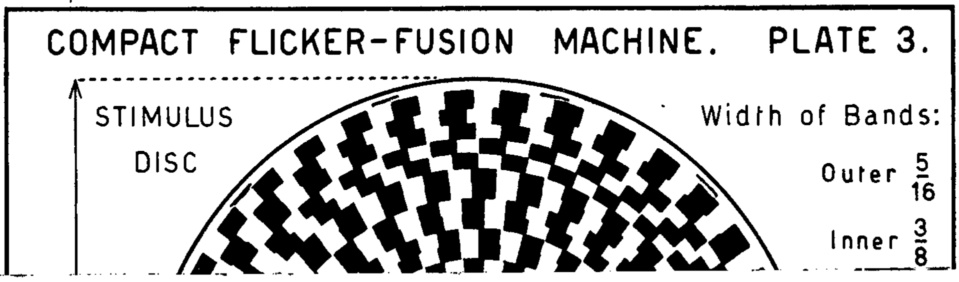
(a) *Mechanical variometer*

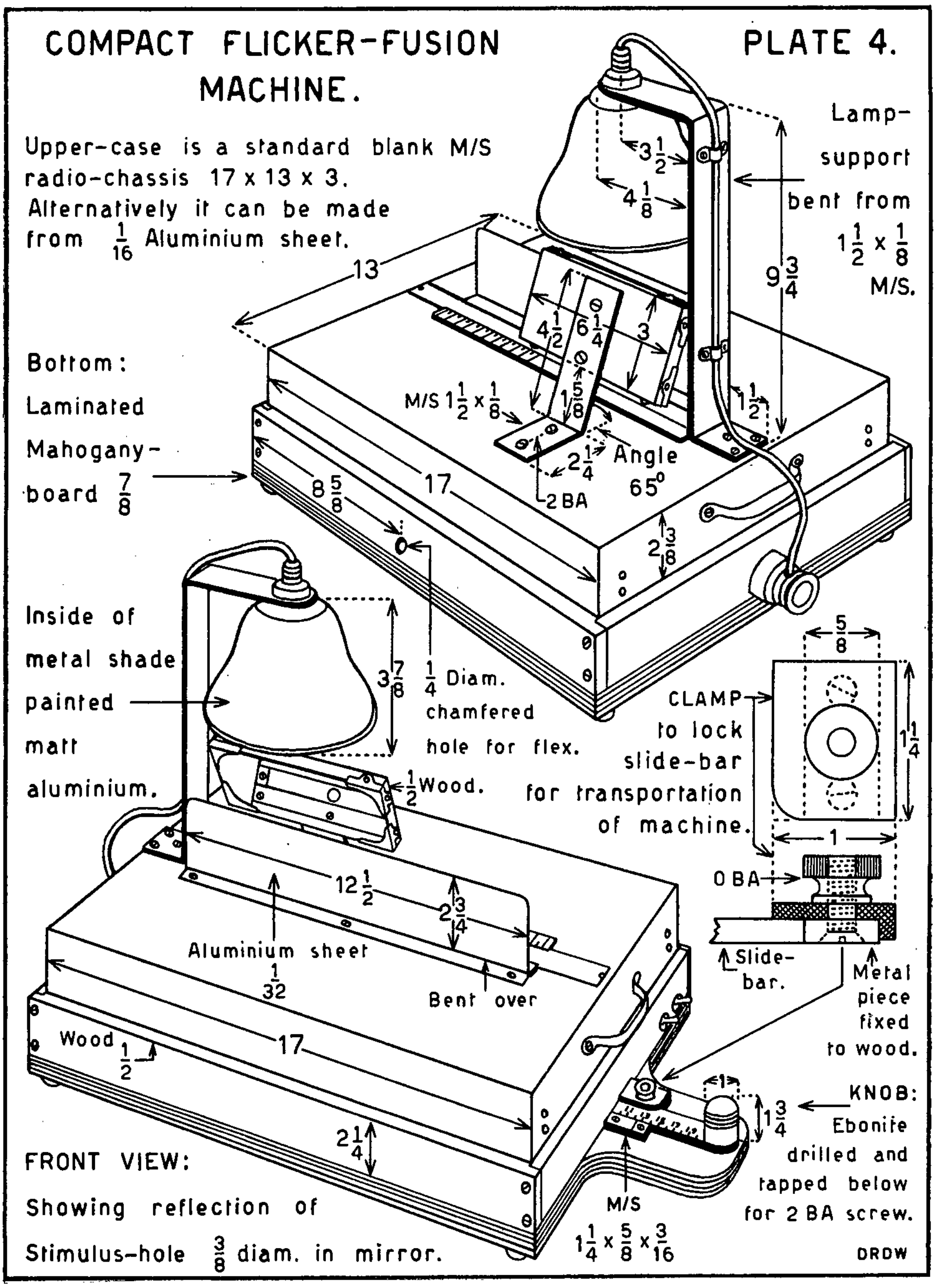
This consisted of a continuously variable mechanical transmission on the friction-drive principle interposed between the constant-speed motor and the rotary

[illegible] ment to permit the speed to be set at a constant 80 r.p.m.

The motor is mounted inside a rectangular sheet-metal box (Plates 2 and 3) so that its spindle for carrying the turn-table protrudes through the top. By means of four flanged rollers at the lower edge of the box, this unit is made to function as a carriage running on two rails screwed to the baseboard inside the wooden chassis of the apparatus. A flat metal slide-bar rigidly attached to the box, projects through a slot in one end of the wooden chassis, and terminating in a vertical finger-knob, allows the entire carriage to be moved to and fro from outside the chassis.

The upper section of the apparatus consists of a rectangular sheet-metal case that rests on a rebate inside the wooden chassis and encloses the carriage and


COMPACT FLICKER-FUSION MACHINE. PLATE 3.
STIMULUS DISC
Width of Bands:
Outer 5/16
Inner 3/8


COMPACT FLICKER-FUSION MACHINE.
PLATE 4.
Upper-case is a standard blank M/S radio-chassis 17 x 13 x 3. Alternatively it can be made from 1/16 Aluminium sheet.
Lamp-support bent from 1 1/2 x 1/8 M/S.
3 1/2
4 1/8
9 3/4
13
4 1/2
6 1/4
3
M/S 1 1/2 x 1/8
1 5/8
1 1/2
Angle 65°
2 1/4
2 BA
Bottom: Laminated Mahogany-board 7/8
8 5/8
17
2 3/8
Inside of metal shade painted matt aluminium.
3 7/8
1/4 Diam. chamfered hole for flex.
1/2 Wood.
CLAMP to lock slide-bar for transportation of machine.
5/8
1 1/4
1
0 BA
Slide-bar.
Metal piece fixed to wood.
12 1/2
2 3/4
Aluminium sheet 1/32
Bent over
Wood 1/2
17
2 1/4
1
1 3/4
KNOB: Ebonite drilled and tapped below for 2 BA screw.
M/S 1 1/4 x 5/8 x 3/16
FRONT VIEW: Showing reflection of Stimulus-hole 3/8 diam. in mirror.
DRDW

turn-table. (Plates 1 and 4.) A longitudinal slot cut in the case-top (Plate 5, Figures A and C) is bounded by two parallel metal strips that extend further along and serve as guides for a masking-strip that can slide in and out to obscure varying amounts of the slot. This masking-strip (Plate 5, Figure B) has a small white dot at one end and, near its middle, a reference-mark that, in conjunction with calibrations engraved on the one guide-strip, indicates the lengths of slot exposed beyond the white dot. Some distance from its other end, the masking-strip has a plain round hole, and when this end of the strip is inserted it provides an alternative and non-variable means of masking the slot by obscuring the entire area except that below the hole.

There is a low sheet-metal screen fitted in front of the slot and guide-strips to prevent the subject from seeing these directly, but a mirror mounted at an angle affords him a reflected view of the slot section. (Plates 1 and 4.)

A lamp-holder and metal shade are supported over the slot-section by a rectangular bracket, and the short lamp-cord ends in a two-pin plug engaging a socket at one end of the chassis. (Plates 2 and 4.) With this coupling disconnected, the upper-case may be lifted off. (Plate 1.)

Electric power is supplied by a flex that enters the chassis on the side away from the subject and passes into a covered junction-block. (Plate 2.) From here distribution takes place to the motor and lamp, each controlled by its own toggle-switch. Enough slack is left in the motor-leads to form a flexible "pig-tail" between carriage and chassis, and so prevent the wire from being broken by the to-and-fro movement.

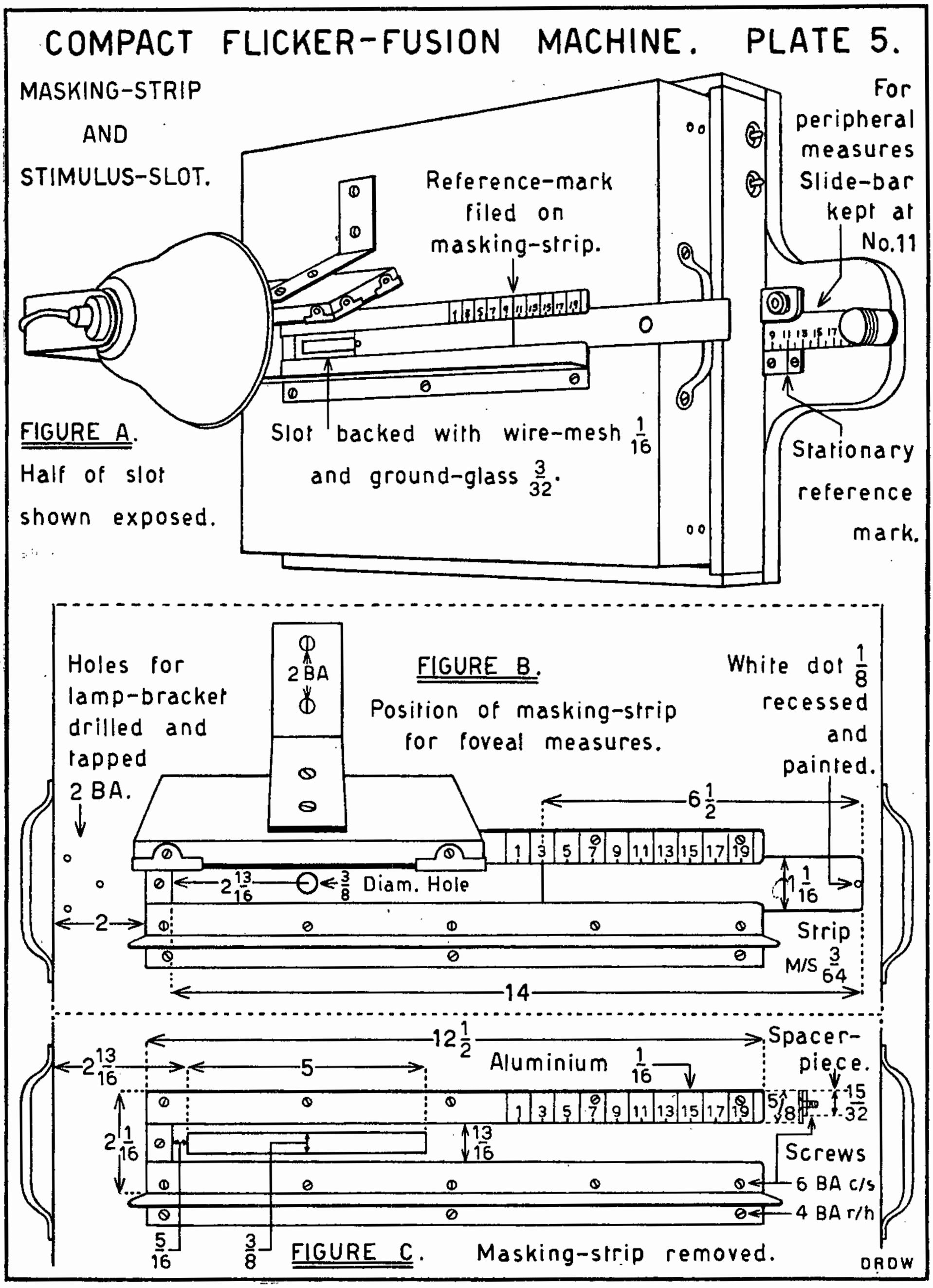
Stimulus-disc

This component, illustrated in Plates 1 and 3, is made from a disc of stiff white card about three sixty-fourths of an inch thick and the full diameter of the turn-table. The stimuli are arranged in twenty concentric bands, and to maintain a standard "one to one" ratio of duration between light and dark phases, each band is divided up into black and white segments of equal size. Proceeding outwards from No. 20 the extreme inner band, the numbers of white (and likewise black) segments in this and the following three, are respectively: ten, fifteen, nineteen, and twenty-two; the next has twenty-three, and for each succeeding band until No. 1, the extreme outer, which has thirty-eight, there is a uniform increase of one segment.

No mechanical dividing-instrument being available, the partitioning of the bands was done with an ordinary pair of draughtsman's dividers set as near as possible to calculated fractions of circumference. Cumulative inaccuracy was avoided by going over the marking several times and making any necessary fine adjustment of the divider-points.

Dark segments are blocked out with matt-black drawing-ink, and a hole punched at the exact centre of the disc fits the tip of the motor-spindle, which has been filed down short and drilled transversely to take a small split-pin. (Plate 2.) At its periphery the disc is secured by wire staples passing through small holes drilled along the rim of the turn-table.

Twenty numbered graduations engraved on the slide-bar controlling the carriage (Plate 2), together with a stationary reference-mark on one guide-block, give settings that refer to the alignment of each stimulus-band directly below


COMPACT FLICKER-FUSION MACHINE. PLATE 5.
MASKING-STRIP AND STIMULUS-SLOT.
For peripheral measures Slide-bar kept at No.11
Reference-mark filed on masking-strip.
FIGURE A.
Half of slot shown exposed.
Slot backed with wire-mesh 1/16 and ground-glass 3/32.
Stationary reference mark.
Holes for lamp-bracket drilled and tapped 2 BA.
2 BA
FIGURE B.
Position of masking-strip for foveal measures.
White dot 1/8 recessed and painted.
6 1/2
2 13/16
3/8 Diam. Hole
1 1/16
2
Strip M/S 3/64
14
12 1/2
Spacer-piece.
2 13/16
5
Aluminium 1/16
15/32
5/8
2 1/16
13/16
Screws
6 BA c/s
4 BA r/h
5/16
3/8
FIGURE C. Masking-strip removed.
DRDW

the round hole in the masking-strip, when this end of the strip covers the slot in the upper-case and the latter rests in position over the disc.

Pulses of reflected light from a miniature "pearl" lamp (40W. 250V. A.C.) are produced at the lowest frequency when the slide-bar is pushed in to graduation No. 20, and at the highest when it is drawn out to No. 1.

To minimize the possibly undesirable effect on the stimulus resulting from inequalities on the disc-surface or slight skewness in the mounting, a diffusing-strip of ground-glass, held by two small clips inside the upper-case, extends the full length of the slot. Interposed between the top of the glass and the slot is a piece of wire fly-screen having a mesh of about one sixteenth of an inch. This reinforces the diffusion without unduly impairing the overall definition of the stimulus, and also fulfils the purpose of aiding the subject to assume his optimum focal distance and keep it during the test.

The smallest unit of measurement, and hence the practical limit of accuracy, is restricted to one photic pulse per three-quarter second, which, although adequate enough for most purposes, might be considered rather crude for fine work. On the total of subjects tested hitherto, the maximum range of scores extended between band No. 3 and band No. 16, inclusive. A stimulus-disc designed to rotate faster than 80 r.p.m., would give smaller units, but this would mean either another type of motor, or additional gearing at the spindle of a standard gramophone-motor.

The use of the reflex principle is a necessary compromise in maintaining compactness and simplicity of design and construction. Reuning (9) found on his apparatus that the most reliable CFF scores were obtained in a darkroom, presumably because of the good contrast between figure and ground. As the principle in the present machine depends on reflected light, there is reduced contrast, both between stimulus and ground, and between light and dark phases in the stimulus itself, which undoubtedly entails some sacrifice of definition.

Testing procedure

By means of a stop-watch and a mark painted on the rim of the turn-table, the speed is adjusted to a constant 80 r.p.m. for all work, and occasional checks are done to ensure that this is being maintained.

The test was administered in ordinary rooms illuminated by slightly subdued daylight, and was so placed that the subject did not face a window. However, controlled artificial light would probably have been more satisfactory.

After the subject has had at least twenty minutes of adaptation to similar indoor lighting conditions, he is seated directly in front of the mirror so that his eyes are approximately fifteen inches away from it.

With the masking-strip obscuring all of the stimulus-slot except that area below the round hole (Plate 5, Figure B), and the slide-bar set so that graduation No. 20 is opposite the stationary reference-mark, the motor and light are switched on, and the subject is asked: "Do you see the flickering in that round spot there?" The reply is invariably in the affirmative, and the subject usually leans forward to assume, what is for him, an optimum viewing distance.

"Don't strain your eyes. Just keep looking at it normally and comfortably. I am going to reduce that flickering, and as soon as you get the impression that it has all stopped and the spot is steady, say 'Yes'."

Using preferably the fingers of both hands on the knob, the observer draws the slide-bar out smoothly at a uniform rate of about one graduation per half-second. Immediately the subject says " Yes," the graduation most nearly opposite the stationary reference-mark is noted and the slide-bar drawn along fairly quickly for its remaining travel to graduation No. 1. On the scoring-sheet, which has two vertical columns of blank spaces, the one labelled " Foveal " and the other " Peripheral," the reading just described is recorded in the first space of the " Foveal " column. The subject is told next: " Say ' Yes ' as soon as the flickering starts again."

The slide-bar is pushed back at the same uniform rate until the subject responds, when the graduation opposite the reference-mark is noted as before and the slide-bar pushed along fairly quickly for the rest of its traverse to graduation No. 20. This reading is recorded in the second " foveal " space, and the subject is then told: " Say ' Yes ' when the flickering stops again." The first and second procedures are alternated in this way until ten " foveal " readings have been obtained.

Both motor and light are now switched off; the masking-strip is withdrawn, and slid back other end first as far as it will go, so that the slot is entirely covered and the white dot is located just beyond its extreme left edge. Graduation No. 11 on the slide-bar is set directly opposite the stationary reference-mark (Plate 5, Figure A), and kept here for the rest of the test. (This position of the carriage puts all bands of stimuli under the slot so that they fill it exactly.)

The motor and light are switched on, and the subject instructed as follows: " Without straining your eyes, just keep looking steadily at that white dot. I shall be moving it along to the right. As soon as you become conscious of flickering anywhere to the left of it where you are not looking directly, but as it were, in the corner of your eye, say ' Yes '."

The masking-strip is then drawn smoothly towards the right, also at a uniform rate of about one graduation per half-second. Immediately the subject responds, the reading is noted, the strip carried along fairly quickly for the remaining distance to the end of the scale, and the reading recorded in the first space of the " Peripheral " column. Then the subject is told: " Just keep looking straight at the white dot, and say ' Yes ' as soon as you are no longer aware of any flickering on the left."

The strip is pushed back at the same rate until the subject responds, when the reading is noted, the sliding-action " followed through " as before, and the reading recorded in the second " peripheral " space on the scoring-sheet. These procedures are repeated alternately until ten " peripheral " readings have been obtained.

Application and discussion

The test was administered experimentally to two groups of pupil-pilot candidates for the South African Air Force, the first in 1956 and the second in 1959. (Table 1.)

As the differences between both the foveal and peripheral means of the two groups (Table 2) were found to be significant at well beyond the 0·001 per cent level, the two sets of results have been treated separately. The cause for this

Table 1

Derivation of samples

	1956	1959
.. Applied for flying course ..	163	152
.. Did flicker test	161	149
Accepted for training	47	55
Suspended from training (all reasons)	13	22
.. Suspended for poor flying ability	8	16
.. Did flicker test	7	16
.. Completed course successfully	33	33
.. Did flicker test	33	31

difference is unknown, but it should be mentioned that the groups were tested in two separate localities, and, amongst other things, conditions of ambient illumination in the respective experimental rooms were probably not identical.

Table 2

Mean critical fusion-frequencies of the two groups

Groups	1956		1959	
	Foveal	Peripheral	Foveal	Peripheral
Total of ten trials	88	36	106	61
S.D.	20·26	11·91	17·04	17·10
Nearest equivalent stimulus-band No. ..	9	4	11	6
Approximate number of photic pulses per second	40	47	37	44
Number of cases	161		149	

For both groups in Table 2, the CFF is higher in peripheral than foveal vision, which conforms to usual findings when the stimulus-area is adequate enough to be easily perceived. In the case of a very small stimulus-area, the opposite result would be expected (Granit and Harper, 3).

Table 3 gives some indication of test reliability by split-half correlations on the "odd-even" pattern between (Trials $1 + 4 + 5 + 8 + 9$) and (Trials $2 + 3 + 6 + 7 + 10$) with corrections by the Spearman-Brown formula for prediction. It is not unlikely that these have been boosted through the "expectancy" element generally inherent in the method of limits, as Reuning (9) has pointed out with reference to other apparently high CFF reliabilities obtained when this technique of administration was used.

Table 3

Product-moment inter-trial correlations

Group	1956	1959
No. of cases	161	149
Foveal r	0·955	0·943
Corrected	0·977	0·971
Peripheral r	0·849	0·926
Corrected	0·919	0·962

For both groups, it would seem that foveal measurement is rather more reliable than peripheral; by rough tests the difference appears to be significant at better than the 5 per cent level. During the peripheral trials, subjects were required to look directly at one point while concentrating their attention on another, and the more noticeable occurrence of indecisive responses in this part of the test, suggests that occasional subjects were unable to resist the temptation of sometimes glancing directly at the stimulus, under the erroneous impression that they would thereby perceive the flicker more easily.

The product-moment correlations between the aggregates of the ten foveal trials and the ten peripheral trials are given in Table 4. They seem large enough to imply that foveal and peripheral CFF are modes of some common function. Had the reliability of peripheral measures equalled the foveal, these correlations might well have been larger.

Table 4

Product-moment correlations between Foveal and Peripheral measures

1956 group ..	0·669	N=161
1959 group ..	0·792	N=149

Correlations obtained between foveal CFF and certain other test variables for the 1956 group, are shown in Table 5. Span of Attention was measured by exposing stationary letters and numbers at a constant shutter-speed of about one-twenty-fifth of a second in a fall-tachistoscope, and determining the maximum number of items subjects could perceive at such a "single glance." In measuring Speed of Perception, cards with numbers and letters were carried by the moving beam past an opening, and the speed was increased for each succeeding presentation in the series. The score consisted of the total number of items correctly identified by the subject.

The testing of Perseveration (motor) was done on a discrimeter having eight keys marked with geometrical symbols in two colours. In the "X" activity, subjects reacted by pressing the appropriate key when an identical symbol appeared

Table 5

Product-moment correlations between Foveal CFF and other tests in the 1956 group

Span of attention	–0·042
Speed of perception	0·105
Errors on hand-foot reaction test	0·055
Learning-time on hand-foot reaction test ..	–0·062
Motor perseveration (speed)	–0·132
Motor perseveration (errors)	0·011
Mental alertness	–0·021

Number of cases = 161

$r \geqslant 0{\cdot}156$ significant at 5% level.

in a small window. In the "Y" activity they had to press the key marked with the corresponding *shape* of symbol to the one in the window but having a different colour. The perseveration-score was mainly a measure of the hindrance-effect of "X" associations over the "Y" activity.

No significant correlation was obtained between foveal CFF and any of the other variables in Table 5, and rough estimates on the peripheral scores were sufficiently convincing in that these also showed no correlation with the above variables. Hence the Visual CFF does not appear to be related to activities that are notably cognitive, which is in agreement with the results of Colgan (2), who found no appreciable relationship between CFF and "attention, memory, perceptual ability, scholastic aptitude or scholastic achievement."

Table 6 gives the biserial correlations between foveal CFF and certain combined temperament and personality ratings of subjects by three observers independently administering three other tests, namely, Arm-leg Co-ordination, Two-hand Co-ordination, and Steadiness (manual). This sample consists of candidates in the 1956 group who were accepted for training, and is therefore curtailed.

Table 6

Biserial correlations between Foveal CFF and temperament ratings

Ratings	r_{bis}	P (one-tailed)
Interested or Bored	0·13	0·241
Sociable or Reserved	0·26	0·092
Emotional or Unemotional ..	0·51	0·003
Number of Cases = 44		

(Positive correlations inply that high CFF is associated with the display of interest, sociability, and emotion)

The implication of these results is that high CFF tends to be associated with the *observed display* of emotional tone, and low CFF with the absence of such display. This would fit in with Wiersma's findings on the Manics and Melancholics. To a much less significant extent the same tendency seems to hold for the *observed display* of sociability. It should be clearly understood, however, that these correlations in no way justify an assumption that subjects with high CFF necessarily incline to be more " emotional " or " sociable " than those with low CFF. There is simply a suggestion that the former manifest their feelings more readily than the latter.

The biserial correlations between Visual CFF and the Success/Suspension flying criterion, are given in Table 7. Suspensions are through lack of competence in actual flying and exclude certain candidates suspended for various other reasons such as medical, educational, disciplinary, and " at own request."

Table 7

Biserial correlations between Visual CFF and success/suspension in flying training

Groups	1956	1959	Combined	P (two-tailed) of mean r_{bis}
No. of cases	40	47	87	
	r_{bis}	r_{bis}	Mean r_{bis}	
Foveal	0·55	0·47	0·50	0·00003
Peripheral	0·46	0·26	0·35	0·006
Aggregate	0·56	0·38	0·46	0·0002

(*A high CFF tends to be associated with success in flying training*)

It appears that subjects with a high CFF stand a better chance of succeeding on a course of flying training than those with a low CFF, and that the foveal measure has greater significance than the peripheral as a predictor of success.

The inferior reliability and validity of the peripheral measure can to some extent have resulted from the particular method used, in so far as movement of the masking-strip altered the stimulus-area, and, as Granit and Harper (3) have shown that the CFF increases with the area of the flickering field, this introduced quite an important variable. Possibly a better method of obtaining peripheral measures, and one more directly comparable with that used for the foveal, would be the following. Let the stimulus be the fixed round hole in the masking-strip, which remains stationary, all control of presentation being done by moving the slide-bar exactly as in the case of foveal measurement. But instead of the subject looking directly at the reflected stimulus-hole, he would fixate a similar sized mark on a small block or vertical card placed on the upper-case to the right of the mirror and at the same apparent focal distance as the reflected stimulus-hole. Alternatively, the length of mirror might be increased and a piece of card carrying the fixation-mark laid flat on the upper-case at the desired point to the right of the stimulus-hole. In either system, subjects would be required to look straight at the fixation-mark and gauge flicker at the stimulus-hole by oblique perception.

However, the retention of an adjustable masking-system is to be recommended, as it may be useful for other experimental applications requiring changes in the size and format of the stimulus.

If, in agreement with Reuning's (9) view, the method of limits is regarded as objectionable on account of the " expectancy " element, settings for both foveal and peripheral measures can be presented in random order by covering the stimulus with a strip of dark card, making the setting, and then exposing the stimulus again, the subject to state each time whether he perceives flicker or not.

The relation between high CFF and success in flying training is not adequately explained by an hypothesis confined to the superior resolving power and recovery rate of the sensory visual process. Reuning (9) reports a significant correlation between CFF and alpha frequency of the electroencephalogram, and from this, in conjunction with Mundy-Castle's (6) correlation between alpha frequency and primary-secondary function, he concludes that: " the individual flicker threshold really seems to depend on some neuro-physiological factor or factors which are related to differences in alpha frequency and operative in one at least of the basic temperament variables." A significant correlation between CFF and alpha frequency was found again, in a larger sample, by Shoul and Reuning (10).

Supporting evidence that the CFF seems to have some connection with temperament is provided in the present study by the correlation of 0·51 ($p = 0{\cdot}003$) between CFF and the observed tendency to free expression of emotion.

There is a suggestion that the candidates with high CFF, namely, those who might be regarded as " primary " or inclined towards mania, have greater freedom from some sort of temperamental inhibition that handicaps the low CFF or " secondary " ones in flying training.

A neat theory to account for the observed differenes in the visual CFF has been suggested by Roberts of this Institute. It is based on the postulate that even small eye-movements accentuate sensitivity for the perception of flicker by distributing the separate photic pulses over a wider field of retinal receptors. He presupposes that however still subjects attempt to keep their eyes in viewing the stimulus, there is always some movement, and it is the varying amount of this movement between individuals that produces the difference in CFF, and not, as is commonly accepted, significant variation in the efficiency of the sensory-visual mechanism. Subjects of a restless or manic disposition exhibit higher CFF than inert ones, because they have more eye-movement, which causes the photic pulses to impinge alternately on different sets of receptors over a wider range.

Such an explanation certainly covers all the known facts, but it requires experimental confirmation that a significant relationship between individual differences in eye-movement and the CFF actually exists.

Shoul and Reuning (10) quote Pfahler's temperament dimension of " wide, fluctuating attention with low perseveration " as characteristic of the primary person. In the course of discussion on present results, Reuning has pointed out that " wide " in the above sense means " wandering," and is not to be confused with " span of attention "; furthermore, that findings from certain experiments mentioned by Pfahler, in which primaries succeeded better than secondaries in distributing their attention over a number of activities at the same time, would offer some explanation for the apparent superiority of the primary in the cockpit of an aircraft, where attention has to be distributed over numerous controls and instruments. In Reuning's opinion, this aspect of "primariness" seems to link up very well with Roberts' theory of eye-movements as a factor in CFF.

Summary and conclusions

(1) After a brief discussion of the general purport of research on the Flicker-fusion threshold (Visual CFF), various systems for its measurement are compared and evaluated.

(2) A full description is given of an apparatus specially designed to be simple, quick, and compact enough for inclusion in a test battery, while conforming, as far as possible, to the following major requirements:

(a) positive and consistent relationship, with minimum time-lag, between the observer's control-settings and changes in the stimulus;
(b) a standard quality of illumination at all frequencies;
(c) freedom from auditory clues betraying changes in frequency;
(d) no fumbling and undue delay between test trials that might disturb the subject's "set";
(e) no necessity for special laboratory accommodation;
(f) testing technique to be either by the method of limits or by random presentation.

(3) Two groups of candidates for acceptance as pupil-pilots in the South African Air Force were given the test experimentally. Reliability, as indicated by inter-trial correlations, appears very high, and has probably been augmented through the use of the method of limits, with its inherent "practice and expectancy" element (Reuning, 9).

(4) For better control of ambient illumination, it might be preferable to conduct the test by subdued artificial light rather than daylight, keeping the source well out of the subject's field of vision, and eschewing any form of discharge-lamp that has perceptible flicker of its own.

(5) Foveal measurement appeared more reliable than peripheral, the difference being significant at better than the 5 per cent. level; to some extent this might have been a consequence of the particular system used in obtaining the peripheral measures.

(6) Product-moment correlations between foveal and peripheral measures seem large enough to imply that these two variables are modes of some common function.

(7) No significant correlations were obtained between the Visual CFF and performance on the following cognitive and sensory-motor tests: Span of Attention, Speed of Perception, Insight-learning and speed on a complex Reaction-test, Perseveration (motor), and Mental Alertness.

(8) A biserial correlation of $0 \cdot 51$ ($p = 0 \cdot 003$) was obtained between high CFF and the tendency to free expression of emotion, as rated by three independent observers in different test situations.

(9) There appears to be a significant relationship between CFF and success or failure on courses of flying training, a high CFF implying a better chance of success. The most valid measurement was the foveal, which gave a mean biserial correlation of $0 \cdot 50$ ($p = 0 \cdot 00003$) with the criterion. It is possible that candidates

having a high CFF are more free from some sort of temperamental inhibition that hampers those with low CFF, and that they have a better faculty for distributing their attention than the latter.

ACKNOWLEDGEMENTS

I wish to thank the following: Miss C. M. Elder, for her kind and efficient assistance with the inter-trial correlations; Mr. J. B. Kirstein, who took the photographs; Dr. L. Kamfer, for the reference to Colgan's work; and Mr. A. O. H. Roberts, for both statistical consultation and most interesting discussion on CFF theory.

REFERENCES

1. BIESHEUVEL, S. *and* D. R. PITT. Some tests of speed and tempo of behaviour as predictors of the primary-secondary function temperament variable. *J. Nat. Inst. Personnel Res.*, 1955, 6, 87–94.
2. COLGAN, C. M. The relation between critical flicker frequency and several psychological variables. Dissertation. *Dissertation Abstracts*, 1954, 14, 1462–63.
3. GRANIT, R. *and* P. HARPER. Comparative studies on the peripheral and central retina. *Amer. J. Physiol.*, 1930, 94, 41–50.
4. GROSS, O. *Die cerebrale Sekundärfunktion.* Leipzig, 1902. (Spearman, p. 43 refers.)
5. LANDIS, C. *An annotated bibliography of flicker fusion phenomena, 1740–1952.* Ann Arbor, University of Michigan, 1953.
6. MUNDY-CASTLE, A. C. The relationship between primary-secondary function and the alpha rythm of the electroencephalogram. *J. Nat. Inst. Personnel Res.*, 1955, 6, 95–102.
7. MURPHY, G. *General psychology,* p. 507. New York, Harper, 1933.
8. PFAHLER, G. *Vererbung als Schicksal. Eine Charakterkunde.* Leipzig, Barth, 1932. (Shoul and Reuning, p. 41 refers.)
9. REUNING, H. A new flicker apparatus for measuring individual differences. *J. Nat. Inst. Personnel Res.*, 1954, 6, 44–54.
10. SHOUL, Shirley M. *and* H. REUNING. Speed and variability components in Pauli test, CFF and alpha rythm. *J. Nat. Inst. Personnel Res.*, 1957, 7, 28–44.
11. SPEARMAN, C. *The abilities of man; their nature and measurement.* London, Macmillan, 1927.
12. WIERSMA, E. Die Sekundärfunktion bei Psychosen. *J. Psychol. Neurol. Leipzig,* 1906, 8, 1–24.
13. WOODWORTH, R. S. *and* H. SCHLOSBERG. *Experimental psychology, rev. ed.* New York, Henry Holt, 1954.

CAPE & TRANSVAAL PRINTERS LIMITED, JOHANNESBURG

SECTION VI

APPARATUS FOR ADMINISTERING COGNITIVE TESTS

Reprinted from " Psychologia Africana," Volume 10, 1964.

A MULTIPLE-CHOICE INSTRUCTIONAL DEVICE

D. R. DE WET

Devices for administering programmed instruction, or " teaching machines " as they are popularly but erroneously designated, are being increasingly applied today as mechanical aids to pupils in teaching themselves the more basic facts or principles in various branches of knowledge.

The method appears to have several advantages over conventional tuition, among these being the emphasis on cultivating self-reliance and responsibility, the immediate and positive reinforcement of correct responses which Skinner (3) and others have stressed as being conducive to greater learning efficiency, the fact that each pupil proceeds at his own optimum rate, so that the fast worker is not retarded by the slow one, nor the latter spurred beyond his capacity, and the removal of much routine drudgery from the manifold responsibilities of the teacher who has more opportunity to give guidance in the synthetic, interpretative, or creative aspects of the work.

A general distinction may be drawn between instructional devices which require a " constructed response ", and those of the " multiple-choice " type (Pressey, 2). Both incorporate the fundamental principle inherent in all programmed instruction, namely, the presentation of the material in a series of short, logically-consecutive units or steps. In the first, the pupil writes his answer to a given item and then compares it with the correct answer; in the second, he is given the choice of several answers and a distinctive reaction by the instrument tells him whether he is right or wrong. The first kind is therefore largely restricted to literate pupils, whereas the second can also be used by illiterates.

The apparatus to be described here is of the latter type (Pressey, 1), but unusual, in that its confirmation mechanism consists essentially of only one moving part—a ball moved by gravity along certain prescribed courses.

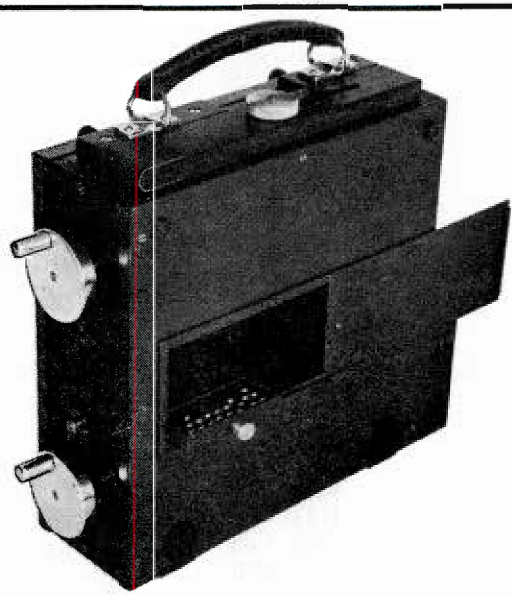
Being non-electrical, it is free from both the shock-hazard always present in mains-operated equipment and the " battery " nuisance of self-contained electrical applicances. Furthermore, it utilizes programmed lessons made up on ordinary, readily-obtainable foolscap paper, on which the material may be written, drawn, typed, or printed with a duplicator. This flexibility has considerable practical significance when such a device is used for research purposes by those interested in compiling their own lessons in preference to applying stock commercial ones, that, as a rule, can only be used in certain proprietary devices intended exclusively for them.

Description

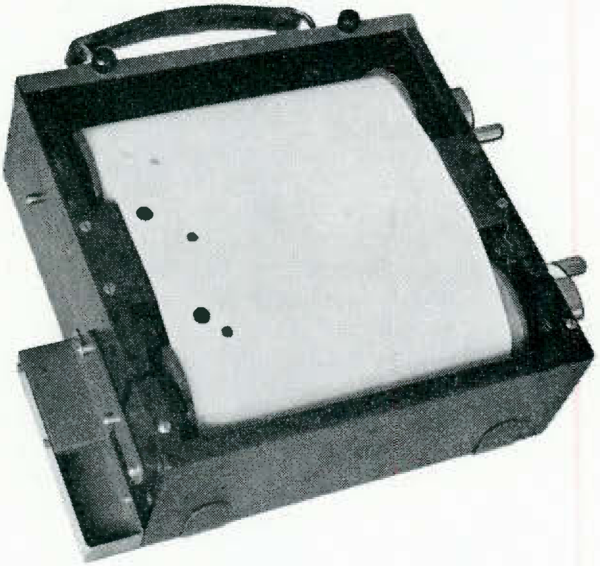
The present machine is rather smaller than an average-sized attaché case, its rectangular body being about ten and a half inches square by three and a half deep. The actual overall dimensions, however, encompassing the various small protruberances, are approximately thirteen and three eighths, by twelve, by five.

Plate 1 (upper figure) and Plate 2 (lower figure) show it set up on the table for use. At the middle of the top, there is a perspex window ($6\frac{1}{2}'' \times 2\frac{3}{8}''$), and at its left end, a small rectangular box perforated with five equidistant holes ($\frac{5}{16}''$ diam.) near the upper left edge. Just to the right of the holes, a card carrying an identification letter or other symbol for each hole reposes beneath a protecting strip of celluloid in a grooved holder fixed to the upper surface of the box. As a lower continuation of the box, there is a perspex-covered tray extending down to the left-hand corner of the case-top.

MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 1.



MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 2.



At the left-hand side of the case, along the lower edge, a metal trough is attached, and on the right-hand side there are two duralumin hand-wheels ($2\frac{1}{4}$ diam. × $\frac{7}{16}''$ wide) fitted with rotating crank-handles.

A wooden strut at the upper end of the case, shown folded away in Plate 1 (lower figure) and extended in Plate 2 (lower figure), supports the case at a forward slope of about thirty degrees in relation to the table on which it stands. The strut has a leather-covered foot and a longitudinal slot along most of its upper half that engages a O B.A. stud projecting through the case just below the carrying handle, and it is clamped in position by means of a substantial knurled finger-nut (1″ diam.) of duralumin. Between the latter and the wood of the strut, a wide, flat washer of thin brass is interposed to prevent the wood from being scored when the nut is tightened. When the machine is to be packed away or transported, the nut is slackened off, the strut swung to one side until it is parallel with the carrying handle, then pushed along the stud the full length of the slot, and the nut tightened again.

Two pads of leather are fixed to the forward lower edge of the case which bears on the table when the machine is set up for use, and four small rubber buffers are screwed to the hardboard bottom, one near each corner (Plate 1, lower figure).

The hand-wheels are fixed, respectively, to the shaft-ends ($\frac{1}{4}''$ diam.) of two rollers inside the case (Plate 3), and a long length of foolscap paper (8″ wide) joined end to end, on which the individual programme-items are typed or otherwise inscribed within frames about the same size as the window, can be wound from one roller to the other on the same basic principle as that employed in the ancient papyrus-book and the modern roll-film camera.

The bearings in the hardwood sides of the case, in which the shaft of the lower roller rests, are slotted (Plate 3), and each is provided with a metal retaining-bar that can be unclamped and hinged back for the removal and insertion of the roller when a programme is changed.

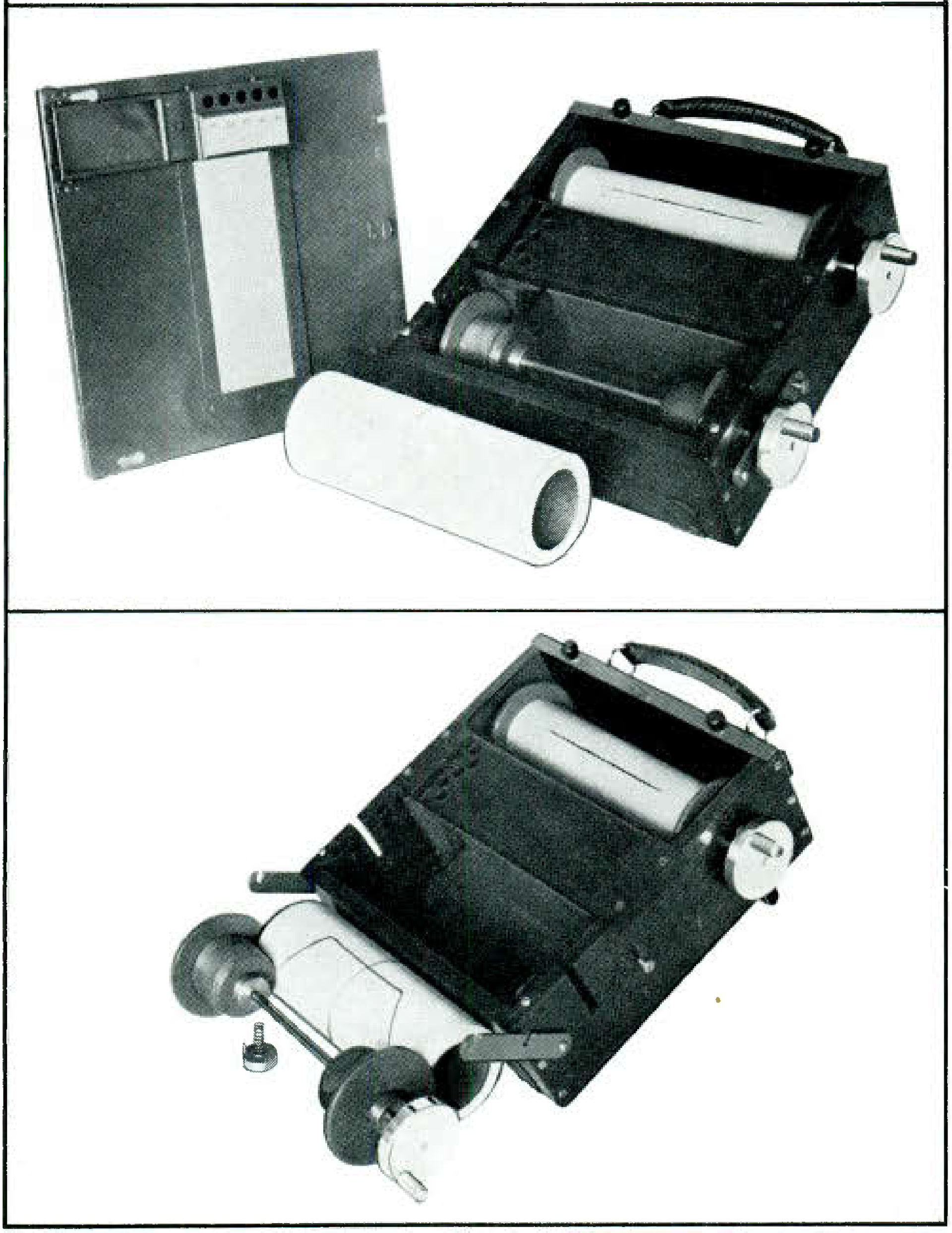
Both retaining-bars are suitably notched to fit over the shaft ($\frac{1}{4}''$ diam.) and the 4 B.A. studs fixed in the wood and on which the free ends are clamped by small knurled finger-nuts.

Each programme is wound on its own cardboard cylinder (8″ × $2\frac{3}{16}''$ O.D.), the end being attached by a strip of adhesive tape, and can be easily slipped in or out of the roller after a small spring-loaded friction-disc and the left-hand bobbin of the roller have been removed. The upper roller, located permanently in its bearings, has a similar friction-disc at its left end. These discs press outwards against the inner faces of the wooden sides and produce a smooth, constant braking-effect that helps to keep the section of paper between the rollers taut against the wooden cross-member under the window.

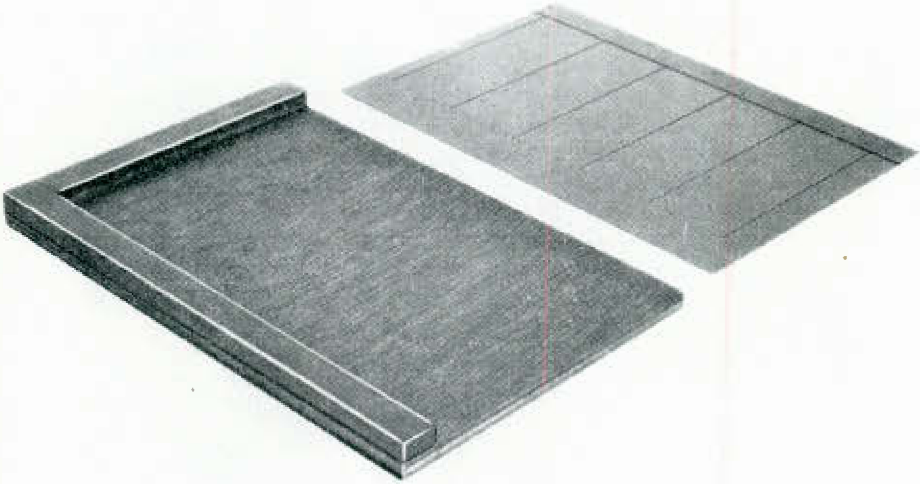
Near its lower end, the removable metal case-top has two " key-holes " that lock under the heads of a pair of screws in the edge of the wood when the top is put on and slid up into position. At the upper end, it is clamped by the two black finger-nuts bearing on a metal angle-strip, and should it be considered advisable to deter the pupil from tampering with the interior, two upright staples midway between the finger-nuts may be temporarily sealed together with thread or other suitable material.

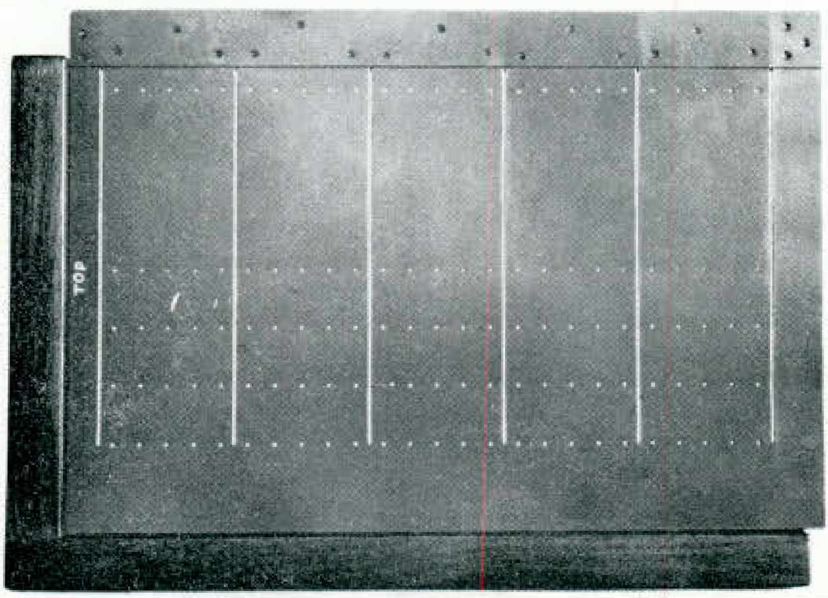
In the floor of the perspex-covered tray, there is a hole ($\frac{5}{8}''$ diam.) kept closed, while the machine is in operation, by a metal tab pivoted on a spring-loaded rivet through the case-top. The hole is opened by moving the free end of the tab to one side, but the tab, being located on the underneath surface of the case-top, can only be reached when the latter is off the case.

MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 3.



MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 4.





Access to a small storage compartment for a shallow, rectangular pan containing a supply of steel balls ($\frac{5}{16}''$ diam.), a few writing materials, and possibly the pupil's lunch-packet, is afforded by a hatchway at the bottom of the case (Plate 1, lower figure). The sliding metal cover has the outer end folded over to give some purchase to the finger-tips, and a screw-and-nut near the inner end that prevents it from being completely withdrawn and very likely lost. When the cover is closed, a spring-loaded button projecting through the hardboard bottom of the case, locks it in place, and provided that the supply-pan has been put like a drawer into its special recess, the case may be turned to any position without dislodging the balls.

The case-top and other external sheet-metal components are made from twenty-two gauge galvanized iron, afterwards spray-painted. Aluminium would produce a lighter structure, but present some difficulty in connection with the soldering of certain inner sections to be mentioned in due course.

Preparation of the programmes

Two accessory pieces of equipment are employed in making up the programmed lessons for use in the machine, namely, the Ruling-plate and board (Plate 4) and the Punching-and-splicing board (Plate 5). One set of each of these will serve a considerable number of machines.

The purpose of the first is simply to provide a convenient means of marking the blank sheets of foolscap paper into the standard areas within which the material of each frame will be inscribed. A frame-width of two and three-eighths inches was chosen because it appeared to be a suitable size for the average length of programmed item and also enabled the most economical use to be made of the foolscap format; five frames to a sheet allowing just the right amount of margin at the top and bottom for pasting the sheets end to end.

Ruling-plate and board

A rectangular piece of laminated mahogany board ($13\frac{7}{8}'' \times 9'' \times \frac{7}{16}''$) has two orthogonal guide-rails of hardwood ($\frac{7}{8}'' \times \frac{1}{2}''$) fixed along the top of one long and one short edge, respectively, as shown in the upper figure in Plate 4.

The Ruling-plate is a rectangular piece of twenty-two gauge mild-steel sheet ($13\frac{3}{8}'' \times 9''$) slotted with six straight saw-cuts ($\frac{3}{64}''$ wide), and perforated with holes ($\frac{1}{16}''$ diam.) arranged in five rows parallel with the long edges. The saw-cuts, of which the first is spaced nine-sixteenths of an inch from the short edge marked " Top ", extend from the one long edge to within one and a half inches of the other, and, to strengthen the plate after cutting, a one-inch strip of the twenty-two gauge material is spot-welded along the edge where the cuts were started.

The first row of holes, which appears as the bottom row in the lower figure in Plate 4, is in line with the ends of the saw-cuts and therefore also one and a half inches from the edge of the plate. The next three rows follow respectively at one-inch intervals, and the fifth is three eighths of an inch from the inner edge of the spot-welded strip. In each row, a group of five holes spaced fifteen thirty-seconds apart, comes within each section demarcated by adjacent saw-cuts, the middle hole of the five being exactly in the middle between the saw-cuts.

The sheets are put on the wooden board, edges against the guide-rails, and the metal plate placed in similar alignment on top. Light pencil lines are then drawn through the slots, and dots stippled through the holes in the fifth or right-hand row to indicate the boundary of the frame, and in the first or left-hand row to indicate the positions to be taken by the five possible " answers " which will appear opposite their respective holes in the perforated box on the machine when the frame is lined up in the window. Only one of the three remaining rows of holes in the plate is stippled, simply to mark the division between the " answer " and " stimulus " areas. If the answers are short, this will be the second or third row, if long, the fourth. It is not strictly necessary to stipple this " dividing " row and the right-hand " boundary " row, entirely. Two dots locating the extremities of each row can be joined by a line drawn on the paper with an ordinary ruler after removing the plate.

Should more copies of a particular programme be required than can be conveniently written or typed and a duplicator must be resorted to, the Ruling-plate can be used for marking the sections on the wax-stencils.

The present prototype Ruling-plate might be improved in the following ways. The three " dividing " rows, and the fifth or right-hand row, could also be in the form of saw-cuts extending between the first and fifth holes in each section, which might speed up the marking of the paper by substituting ten quick pencil-strokes for fifty stipples or four stipples and two additional ruling processes. Alternatively, the above modification could be restricted to the three " dividing " rows, and the width of the plate decreased by three-eighths of an inch, which would enable the fifth or right-hand row to be dispensed with, as the boundary of the item-frame would then be indicated by the ends of the pencil lines drawn to the inner edge of the strip.

Punching-and-splicing board

When the programmed material is on the foolscap sheets, they must be joined together lengthwise to form a long band, and two punch-holes made in the left-hand margin opposite each frame, as shown in the upper figure in Plate 2. If preferred, however, the punching may be done before the splicing.

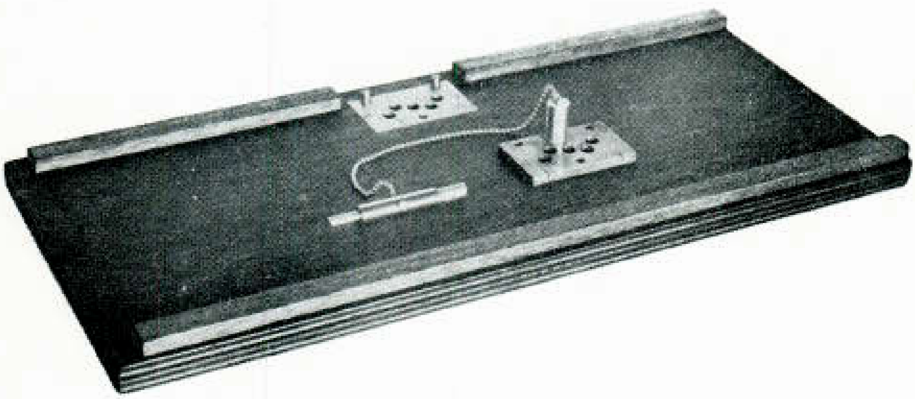
Both processes are done on the apparatus shown in Plate 5, which consists of a rectangular piece of laminated mahogany board ($23'' \times 10'' \times \frac{3}{4}''$) fitted with parallel rails ($\frac{3}{4}'' \times \frac{3}{8}''$) spaced eight and one-eighth inches between their inner edges, and a three-piece punching outfit, of which the lower die is fixed flush in a recess in the surface of the board partly overlapping the edge of one rail where a " cut-away " has been left, while the upper die and punch are removable.

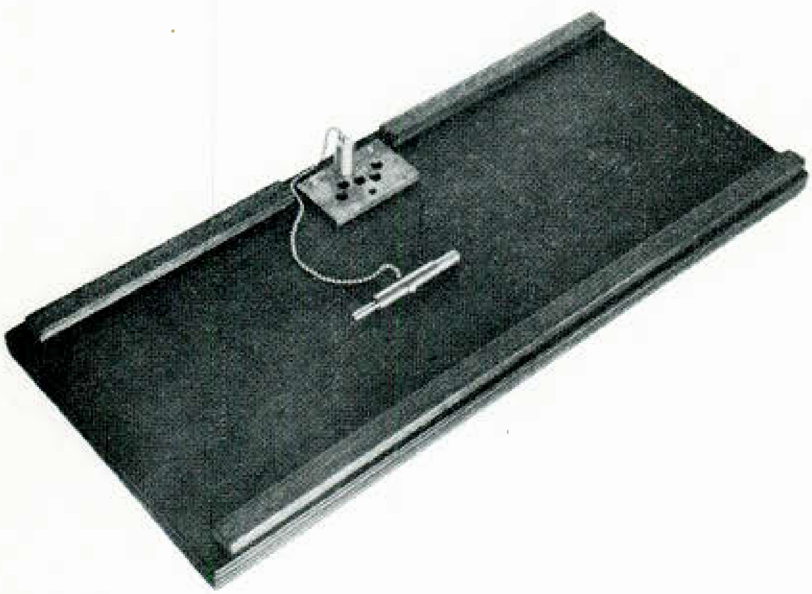
The dies have five holes ($\frac{15}{32}''$ diam.), corresponding in layout to another set of holes ($\frac{25}{64}''$ diam.) near the left-hand end of the wooden cross-member over which the paper passes from one roller to the other (Plates 3 and 8). In both dies, a smaller hole ($\frac{17}{64}''$ diam.) is located five-eighths of an inch to the right of the middle one in the group, where the five holes are positioned equidistantly but in a slightly staggered formation in order to compress their vertical spread well within the limit of two and three-eighths inches allowed by the width of the item-frames.

Both the die-plates are of mild-steel ($3\frac{5}{16}'' \times 2'' \times \frac{1}{4}''$), and when they are together alignment is maintained by two upright pins ($\frac{1}{4}''$ diam.) near the outer edge of the lower which engage registration holes in the upper.

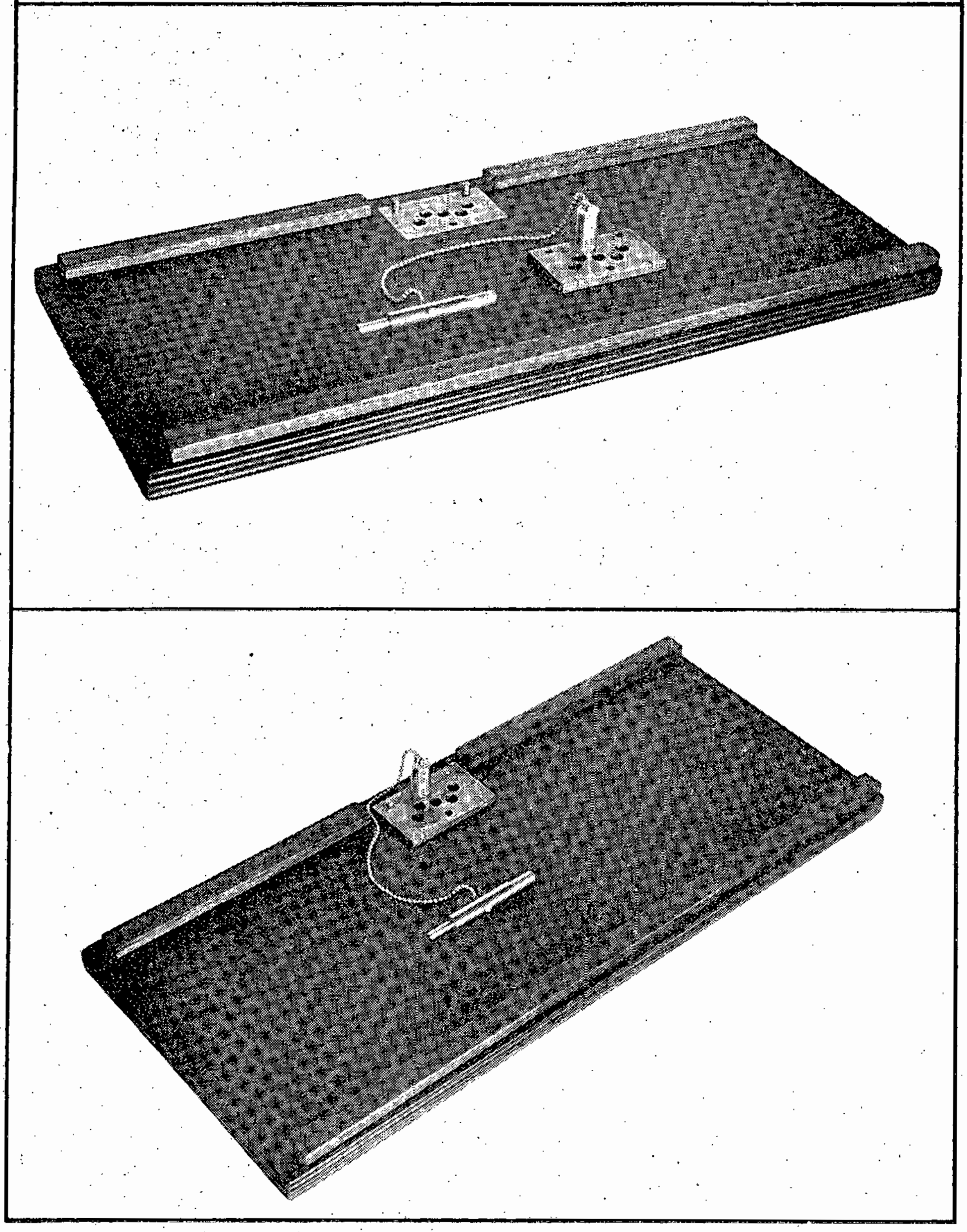
The double-ended punch, attached by a length of light, flexible chain to the short duralumin handle of the upper die-plate, simply to reduce the chance of either component getting mislaid, consists of a duralumin handle ($\frac{3}{4}''$ diam. $\times$ 3″) having a point of silver-steel rod recessed in either end and held by a grub-screw. Both points project

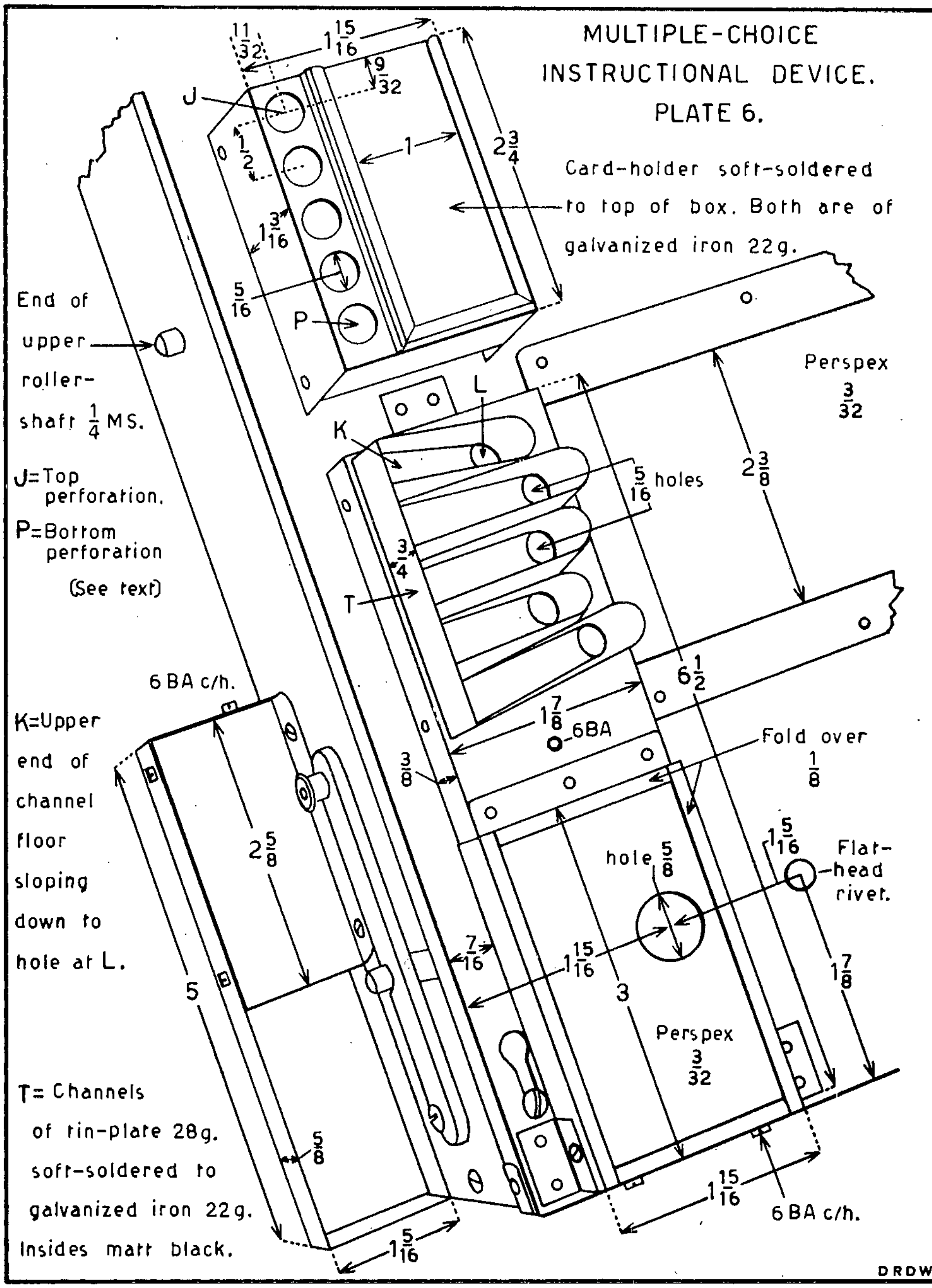
MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 5.





MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 5.




MULTIPLE-CHOICE
INSTRUCTIONAL DEVICE.
PLATE 6.
Card-holder soft-soldered to top of box. Both are of galvanized iron 22g.
End of upper roller-shaft 1/4 MS.
J=Top perforation.
P=Bottom perforation
(See text)
K=Upper end of channel floor sloping down to hole at L.
T= Channels of tin-plate 28g. soft-soldered to galvanized iron 22g. Insides matt black.
Perspex 3/32
5/16 holes
Fold over 1/8
hole 5/8
Flat-head rivet.
6 BA c/h.
6BA
DRDW

seven-eighths of an inch, the diameter of the one being fifteen thirty-seconds to fit the larger holes in the die-plates, and the other, seventeen sixty-fourths, to fit the smaller " indicator " hole. The punch ends are filed concave to provide good cutting edges, and are then hardened. If a great amount of punching is to be done, it is recommended that the die-plates be case-hardened, or made instead, from quarter-inch gauge-plate, heat-treated in the usual way.

Before the punching is started, it is worthwhile to mark by a pencilled letter or number at the extreme left-hand margin of each frame, which of the five holes must be punched for the correct " answer ". Like the hole itself, such a mark should be invisible to the pupil when the programme is in the machine.

The paper sheets are put between the dies, each frame lined up in turn with a pair of saw-cuts at the edge of the upper die, one " indicator " perforation made by inserting the thin end of the punch through the small hole, and one " correct answer " perforation (through which the balls are intended to drop) by inserting the thick end through the appropriate one of the larger holes. The guide-rails keep the paper in alignment, and as each frame is dealt with, the upper die-plate is raised slightly to enable the sheet to be slid along.

As there is often a slight variation in the size of paper sheets from different manufactured batches, it is advisable, when doing the punching, to regard the guide-rail on the punch side as the " master " rail, and always keep the edge of the paper up against it, irrespective of whether the other side of the paper touches the other rail or not. The same precaution should be taken during the splicing process, to be described next. This ensures that the holes in the paper are always at their proper distances from the left-hand margin, which is important if they are to coincide properly with their respective channels on the machine.

In the construction of the punch and die-plates, the registration between the plates and the fit between punch-ends and holes should be fairly accurate if clean, sharp-edged holes are to be produced in the paper.

Having removed the upper die and punch from the board, the latter is used as a guide in joining the sheets end to end. Gum is applied for a distance of about half an inch all along the edge of one sheet located at the end of the board where the guide-rails are stepped back ($\frac{5}{8}''$), and after moving the gummed part well up between the rails, the edge of the second sheet is pressed on to form a half-inch overlap. The two joined sheets are then moved up along the board, gum applied to the end of the second, and the procedure repeated. When about six sheets (of which the first couple should be blanks to provide a " leader ") have been joined, it is convenient to trim the end to a taper, insert this in the slit of the upper roller in the machine (Plate 3) after the latter has been placed on the table in line with the board, and then wind the programme up as the remaining sheets are pasted on.

Application

To bring the first frame into view in the window, the pupil turns the hand-wheel furthest from him in a clockwise direction until the " indicator " punch-hole in the paper is in line with a small white reference mark at the middle of the left-hand edge of the window (Plate 7). He then inspects the " question " in the frame and essays an " answer " by taking a steel ball ($\frac{5}{16}''$ diam.) from the loose supply-pan that rests on the table within easy reach, and drops it through what he regards as the " correct " perforation in the box.

If his choice *is* correct, the ball falls through the margin of the paper via the larger punch-hole which is not visible to him, continues through the corresponding hole in the wooden cross-member (Plates 3 and 8) into a metal " gathering-trough " below, and is ejected through a chute and an orifice near the lower edge of the case (Plate 9) into the " delivery-trough " outside at the left (Plates 1 and 2), where it again becomes available for use on the next programme-item, which is brought into view and aligned by turning the upper hand-wheel again to wind on the paper until the next " indicator " punch-hole is directly opposite the white reference mark.

If the pupil's answer is incorrect, the ball, having been dropped through the wrong perforation in the box, strikes an unpunched area of the paper margin and is deflected by one of a set of fixed baffles on the case-top under the box, into the perspex-covered tray at the corner of the case-top, where it remains visible but is forfeited as a token of error. Should this occur, he again takes a ball from the supply-pan and tries another answer by dropping it into some other perforation in the box, repeating this until confirmation of the correct answer is obtained when a ball emerges in the delivery-trough.

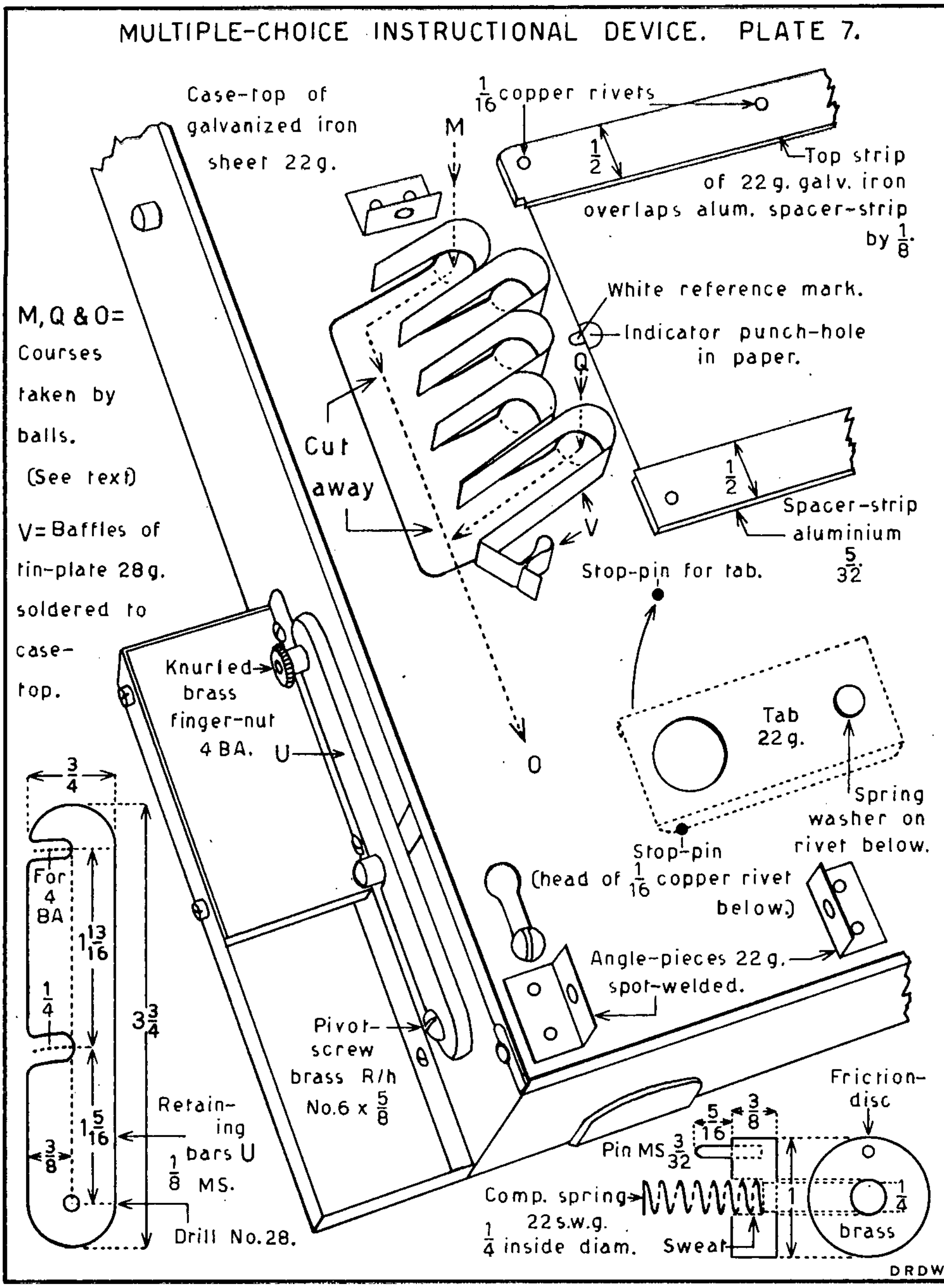
In Plate 1 (upper figure), the item or " question " in the frame happens to be a capital " A ", occurring in a programme of a very elementary sort in which illiterate pupils are required to identify certain letters of the alphabet, presented randomly, one at a time, in script of varying size and kind. For this item, if a ball is dropped in the perforation marked " A " on the indicator-card alongside, it rolls through into the delivery-trough and thereby confirms that the response is correct. If it is dropped into any one of the others marked B, C, D, and E, respectively, it rolls into the perspex-covered tray to indicate and record an error.

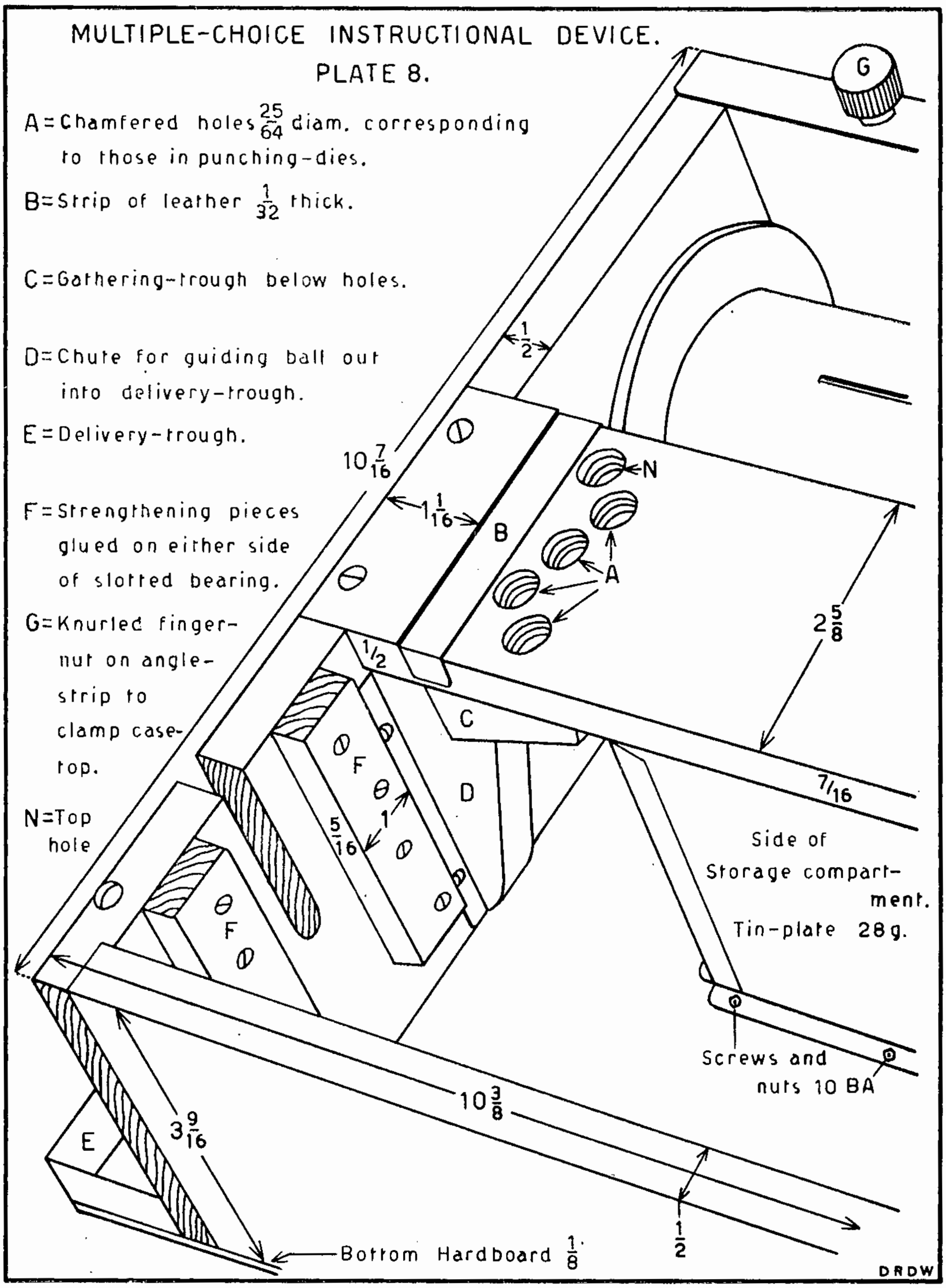
It is important that the " indicator " punch-hole be in line with the white reference mark, otherwise the larger punch-hole may not be properly in register with its channel, and the ball, instead of dropping through to indicate a correct response, may strike the edge of the punch-hole and roll off the paper as an error. If the " indicator " punch-hole is inadvertently wound beyond the mark, it is easily brought back into line by moving the other hand-wheel the required amount in the opposite direction.

The author has been informed by Dr. H. L. Krige of this Institute that children in the first-year primary school were found to have no difficulty in turning accurately on to the mark each time. Sometimes, however, in their eagerness to insert the balls, they fumbled with them and dropped them on the floor. Increased familiarity with the machine would probably reduce this to a negligible amount.

When the programme is of a rather more advanced sort, in which, for example, the illustration of a cat appears in the frame together with the five possible answers Dog; Cat, Pig, Hare and Rat, from among which the pupil has to choose the correct one, the stimulus-item is arranged at the right side of the frame and the answers in a vertical column at the left, so that each one will be directly opposite a perforation in the box, in which case, an indicator strip on the box is not required. This system would also be applied to programmes of complex material, such as mathematical equations.

When the pupil has completed all the items on the programme, and almost all the paper has been wound from the lower roller to the upper, the word " Stop " or some other equivalent sign appears in the window, and the programme is then rapidly rewound on to the lower roller preparatory to removal. After taking off the case-top, the tab closing the hole in the floor of the perspex-covered tray is pivoted to one side, and the " error " balls are tipped out into the supply-pan for use in the next lesson.


MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 7.
Case-top of galvanized iron sheet 22g.
1/16 copper rivets
M
1/2
Top strip of 22g. galv. iron overlaps alum. spacer-strip by 1/8.
White reference mark.
Indicator punch-hole in paper.
Q
M, Q & O= Courses taken by balls. (See text)
Cut away
1/2
Spacer-strip aluminium 5/32
V
V= Baffles of tin-plate 28g. soldered to case-top.
Stop-pin for tab.
Knurled brass finger-nut 4 BA.
U
O
Tab 22g.
Spring washer on rivet below.
Stop-pin (head of 1/16 copper rivet below.)
3/4
For 4 BA
1 13/16
1/4
3 3/4
Angle-pieces 22g. spot-welded.
Pivot-screw brass R/h No.6 x 5/8
Friction-disc
Retaining bars U 1/8 MS.
3/8
1 5/16
5/16
3/8
Pin MS 3/32
Drill No.28.
Comp. spring 22 s.w.g. 1/4 inside diam.
Sweat
1
1/4
brass
DRDW


MULTIPLE-CHOICE INSTRUCTIONAL DEVICE.
PLATE 8.
A = Chamfered holes 25/64 diam. corresponding to those in punching-dies.
B = Strip of leather 1/32 thick.
C = Gathering-trough below holes.
D = Chute for guiding ball out into delivery-trough.
E = Delivery-trough.
F = Strengthening pieces glued on either side of slotted bearing.
G = Knurled finger-nut on angle-strip to clamp case-top.
N = Top hole
G
N
A
B
C
D
E
F
F
10 7/16
1 1/16
1/2
2 5/8
1/2
5/16
1
7/16
Side of Storage compartment. Tin-plate 28 g.
Screws and nuts 10 BA
10 3/8
3 9/16
1/2
Bottom Hardboard 1/8
DRDW

Sometimes it might be advisable for literate pupils to record each correct answer and the number and sequence of any errors made in obtaining it, on a paper form separate from the machine. They should also mark those correct answers which they do not understand and would like to discuss with the teacher.

The machine can perform the additional function of administering short periodic tests to sound the pupil's progress. It can also be used for applying a variety of psychological tests intended to measure the ability to derive systematically coded relationships between certain configurations.

Details of the confirmation-mechanism

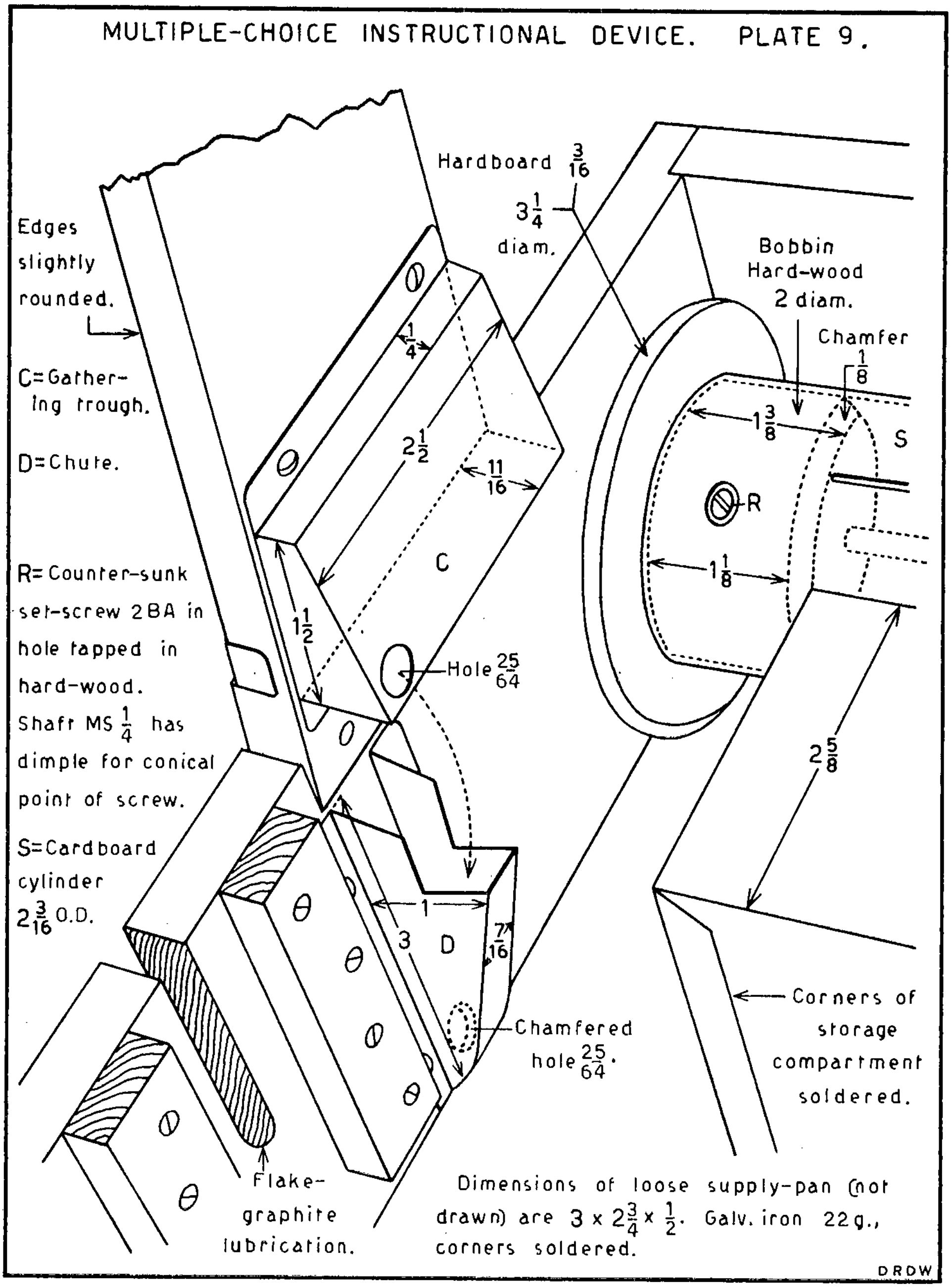
As shown in Plate 2 (upper figure) and Plate 3, the paper passes from one roller to the other over a wooden cross-member bridging the left and right sides of the case and having its ends recessed flush with their upper edges. The top surface of this cross-member is relieved to a depth of one-sixteenth of an inch to provide an easy clearance for the taut paper when the metal case-top is in position, and near the left-hand end are the five holes ($\frac{25}{64}''$ diam.) corresponding in both their relative arrangement and distance from the left-hand paper-margin, to the five large holes in the punching-dies. The distances from the margin are: $\frac{1}{2}''$ to centres of outer two, $\frac{11}{16}''$ to centre of middle one, and $\frac{7}{8}''$ to centres of inner two. Vertically, the top and bottom holes have their centres $\frac{7}{8}''$ from the centre of the middle one, and the intermediate holes are located symmetrically in between. A close-up view of this section is given in Plate 8. Just to the left of the holes, a strip of leather ($\frac{1}{32}''$ thick) glued to the wood serves to keep the edge of the paper up near the underneath side of the metal case-top when it is in position. Alternatively, a low wooden step may be left here.

Screwed to the lower surface of the cross-member, below the holes, is the sheet-metal " gathering-trough " having a sloping bottom perforated at the lower corner with a hole ($\frac{25}{64}''$ diam.) communicating with a metal chute attached to the inside of the case. At the lowest point of the chute, a chamfered orifice leads through the side of the case into the delivery-trough outside. In Plate 9, the cross-member is shown raised at one end to display the arrangement below.

In Plate 6, the case-top is in position, but the outer shell of the perforated box has been taken off to show a group of five sloping channels of twenty-eight gauge tin-plate, which terminate, respectively, at the floor of the box, in holes directly over their counter-parts through the case-top and the cross-member below.

The object in having these channels, instead of simply locating the perforations in the box-top directly over the other sets of holes, is to prevent the pupil from perceiving the " correct answer " punch-hole in the paper merely by looking straight down the perforations. If this were permitted, the main practical value of the machine as an instructional device would be lost.

In Plate 7, the channel-section of the box and the perspex-covered tray, which is just an extension of the same structure, have been removed to show the openings in the case-top. It will be seen that the holes are slotted towards the left; each is


MULTIPLE-CHOICE INSTRUCTIONAL DEVICE. PLATE 9.
Hardboard 3/16
3 1/4 diam.
Bobbin Hard-wood 2 diam.
Chamfer 1/8
Edges slightly rounded.
C=Gathering trough.
D=Chute.
R=Counter-sunk set-screw 2BA in hole tapped in hard-wood. Shaft MS 1/4 has dimple for conical point of screw.
S=Cardboard cylinder 2 3/16 O.D.
1/4
2 1/2
11/16
C
1 1/2
Hole 25/64
1 3/8
1 1/8
R
S
2 5/8
1
3
D
7/16
Chamfered hole 25/64.
Corners of storage compartment soldered.
Flake-graphite lubrication.
Dimensions of loose supply-pan (not drawn) are 3 x 2 3/4 x 1/2. Galv. iron 22g., corners soldered.
DRDW

provided with a sloping tin-plate baffle open at the left; and the slots merge into a common " cut-away " area in the case-top at the ends of the baffles.

Referring to Plate 6 again, suppose that the pupil has decided that the correct " answer " for the particular frame in the window is the top perforation " J " in the box, and accordingly drops the ball into it. The ball strikes the raised end of the corresponding channel at " K ", rolls down and falls through the hole " L ". Referring now to Plate 7, if there is a hole punched in the paper at " M ", his answer is correct, and the ball then drops through hole " N " (Plate 8), strikes the sloping bottom of the gathering-trough " C " (Plates 8 and 9), rolls down along the lower or left-hand edge of the gathering-trough, and, as shown by the dotted arrow in Plate 9, drops into the chute and is ejected into the delivery-trough " E " (Plate 8). However, if the paper is unperforated at " M " (Plate 7), the ball strikes its surface, rolls along the side of the baffle, and is deflected out into the " error " tray at " O " (Plate 7). Similarly, if the ball is dropped into the bottom perforation in the box at " P " (Plate 6), it reaches position " Q " (Plate 7), and depending on whether the paper has a punch-hole here or not, takes one or other of the above courses. The same principle applies to the other channels. Either event takes place within a couple of seconds.

ACKNOWLEDGMENT

I wish to thank Dr. H. L. Krige for his helpful comments on the manuscript.

REFERENCES

1. PRESSEY, S. L. A simple apparatus which gives tests and scores—and teaches. *Sch. Soc.*, 1926, 23, 373–376.

2. PRESSEY, S. L. Teaching machine (and learning theory) crisis. *J. appl. Psychol.*, 1963, 47, 1–6.

3. SKINNER, B. F. The science of learning and the art of teaching. *Harv. educ. Rev.*, 1954, 24, 86–97.

Mr. D. R. de Wet,
National Institute for Personnel Research,
P.O. Box 10319,
Johannesburg,
Republic of South Africa.

Manuscript received: 1st April 1964.
Revised manuscript received: 1st May, 1964.

PRINTED BY CAPE & TRANSVAAL PRINTERS LIMITED, JOHANNESBURG

Reprinted from "Psychologia Africana," Volume 11, No. 2, 1966.

PAIRED-ASSOCIATE LEARNING THROUGH A PRACTICAL MEDIUM

D. R. DE WET

Using an instructional device of the multiple-choice type, a test of paired-associate learning was administered experimentally to 177 air-pilot candidates. Results indicate that it has some potential usefulness for the prediction of success in certain branches of the flying-course, particularly, practical airmanship. However, like most learning tests it requires a rather long time to administer as compared with tests where the emphasis is more on initial ability, and further study will be necessary to determine whether its inclusion in a standard selection battery is expedient or not. It showed small to moderate correlations with certain other tests, intellectual perceptual and sensory-motor, but none with temperament and personality variates of the introversion/extraversion kind.

Among the many exacting tasks which have to be mastered by the pupil air-pilot, is the acquisition and implementation of correct associations between sense data and responses. Certain things, such as instrument-readings, control-positions, or combinations of these, have a definite mutual significance that must be learned and utilized by the pilot to make his aircraft behave most satisfactorily under various circumstances. Although instruction both in theoretical and practical airmanship emphasizes these important relationships, much of the learning still depends on trial and error by the pupil himself. This conclusion is supported by validity coefficients obtained between flying success and the learning section on the rather complex Hand-foot reaction test (1).

A constructionally simpler test of trial and error learning, on the paired associate principle, involving abstract elements in a practical situation where respective right and wrong associations produced two markedly different physical results, was administered to a group of air-pilot candidates by means of the Multiple-choice instructional device (2). The main objects were to investigate: (a) Its validity for predicting success on the flying-course; (b) Correlations between measures and between sections within the test; and (c) Relationships to various other tests, intellectual, perceptual, sensory-motor and temperamental.

Material and Procedure

The learning programme, on a continuous paper roll, consisted of three sections of twenty frames each. In the first, a capital letter A, B, C, D or E was inscribed in each frame, the initial four being in alphabetical sequence, and the rest in random sequence. In the second, each frame contained a numeral 1, 2, 3, 4 or 5, and in the third, a square figure with a small ring located inside one of the four corners or at the middle. In sections two and three the sequence was random from the initial frame. Spaces, preceding and terminating each section, contained respectively, the words " Begin " and " Stop ".

As described elsewhere (2), the left-hand margin of the paper was punched with small holes to facilitate frame-alignment in the window of the apparatus, and larger holes to pair each frame with an appropriate channel in the apparatus, care being taken to ensure that the channels were paired randomly to the stimulus-items to avoid having any obvious association between the positional order of the channels and the normal sequence of letters and numerals or the position of the small ring in the square.

A strip of card inscribed with the following symbols:

was placed in the holder alongside the five holes in the box on the apparatus, each symbol being directly in line with a hole. The subject was seated comfortably before the apparatus and given the following instructions:

> " These five symbols stand for the ' A, B, C, D and E ' in a certain code. Your task is to find out which one is ' A ', and ' B ' and so on. You turn this wheel to move the paper along until the round hole is directly opposite that little white mark. (Demonstration.) The letter in the window is ' A '. To find out which of the symbols it corresponds to, drop a marble into the hole next to one of the symbols. If the marble rolls out here, you are right and have discovered the symbol for ' A '. But if it rolls into this box, you are wrong and must take another marble and try one of the other symbols, repeating this until the marble rolls out and you are right.
>
> You then turn on the paper to bring the next letter into view, make sure that the round hole is directly opposite the little white mark, and drop in a marble or more to find out what symbol the second letter corresponds to. As soon as you have the right answer, turn on to the next letter, and always remember that the round hole in the paper *must* be in line with the white mark before you drop in the marble, otherwise it may not roll out, even when you *are* correct. If you should happen to turn the paper on a little too far, just bring it back in line with this lower wheel. (Demonstration.)
>
> Work as quickly as you can, and try to make the smallest number of mistakes in identifying the five symbols. Try to remember exactly what each symbol stands for because later you will have to identify them somewhere else."

The tester turns the paper back to " Begin ", tells the subject to proceed, and starts a stop-clock. At the end of the first section, the time taken by the subject and the number of marbles in the " error " box is recorded.

For the second section, the subject is told: " Each of these symbols has a certain value, 1, 2, 3, 4 or 5. Proceed in the same way to find out the value of each. Work as quickly as you can, with the fewest mistakes, and try to remember their values." After time and error has been recorded, the third section follows:

> " Each of these symbols also represents a position in a square marked by a small ring, which will be in one of the corners or at the middle of the square. Find out the relationship as quickly as you can, with the fewest mistakes, and try to remember it."

When time and error has been recorded, the subject is handed a small slip with the five symbols printed along the top and three blank rows below, and told: " You have two minutes in which to fill in here, below each symbol, the corresponding letter, value, and position of the small ring in the square as you have learnt them. You may start and finish anywhere you like." The symbol-strip in the holder on the apparatus is removed from the subject's view.

As a precaution against possible " coaching " of other subjects in the waiting-room by those who had finished, after testing each subject the symbol-strip was changed for another with the symbols in a different order, four sets being used in rotation.

Results and Discussion

Intercorrelations between the three sections of the test on the respective measures Time and Number of errors are given in Table 1. All the coefficients are significant at better than the 1 per cent level, but those for " Errors " are substantially smaller than those for " Time ". This suggests that in a learning task of this kind errors provide a stronger reflection of elements of specificity in the material of the different sections than does the rate of work, which has more individual constancy throughout the three sections. It has been reported by Howie (3) that apart from subjects' ability in various tests, there is evidence of a broad speed factor " indicating an individual difference trait of tempo in functioning ".

Table 1

Intercorrelations between sections of test (N = 177)

Sections	Time	Errors
1, Letters and 2, Numbers	0.47	0.26
1, Letters and 3, Ring position	0.46	0.29
2, Numbers and 3, Ring position	0.47	0.20

All significant at better than the 1 per cent level.

As shown in Table 2, the correlation between " Time " and " Errors " in all three sections is large, mainly because the particular test-procedure used inevitably introduced some dependence between these measures. An item appearing in the window had to be correctly identified before the subject proceeded to the next, and each attempt at identification by dropping in a marble took additional time, hence the greater the number of mistakes made, the slower the progress. The relationship, however, was not reciprocal, for, while the making of errors added to the time, working at a slower rate did not necessarily produce more errors. A more independent measure of rate of work would have been obtained if subjects had been allowed only one attempt at each stimulus item as was done on the Hand-foot reaction test (1), but this would have made the learning task much more difficult and imposed the necessity for a longer testing session than was practicable under the circumstances.

Table 2

Correlations between Time and Errors (N = 177)

Sections	r
1, Letters	0·57
2, Numbers	0·61
3, Ring position	0·49
Total 1 + 2 + 3	0·37

All significant at better than the 1 per cent level.

Referring to Table 3, it will be seen that differences in the mean time taken to complete each section do not reach the 5 per cent level of significance, which requires a critical ratio of at least 1·95. On the other hand, differences in the mean number of errors made in each section are very significant, and there is a progressive increase in the means from the first section to the third. It was expected that the general effect of practice at manipulating the apparatus and establishing associations between symbols and stimulus-items in an earlier section would facilitate the learning process in a subsequent section, but the results indicate just the opposite. It is possible that the differences relate more to the nature of the material in the respective sections than the experimental sequence of the sections; alphabetical letters having greater associative affinity with the reference symbols, than numbers, and ring-positions, less. The latter section, where most errors were made, differs notably from the other two in that the cue for item-recognition is spatial rather than configurational. However, the influence of experimental sequence cannot be ignored. This probably induced a certain amount of negative transfer or interference whereby the establishment of associations between the symbols and the stimulus-items in a preceding section tended to impair the acquisition of a new set in a subsequent one.

Retention

A similar effect could have persisted in the last part of the test where subjects had to recall what they had learnt. The means indicate that significantly more of the associations in the first section were recalled correctly, than of those in the second or third section. But here, too, there is the likelihood that the strongest associations were between alphabetical letters and the reference symbols.

Table 3

Means, Standard Deviations and Critical Ratios of the differences between means ($N = 177$)

Measure	Section	Mean	S.D.	C.R. of Diff. M's	
TIME (Seconds)	1, Letters ..	211·06	60·21	1 & 2	0 85
	2, Numbers ..	207·05	61·78	1 & 3	1·03
	3, Ring position..	216·08	63·79	2 & 3	1 85
ERRORS	1, Letters ..	12·40	9·47	1 & 2	5·90
	2, Numbers ..	17·91	10·75	1 & 3	10·39
	3, Ring position..	22·03	11·08	2 & 3	3·92
RECALL (Maximum 5)	1, Letters ..	2·59	1·62	1 & 2	5·48
	2, Numbers ..	1·91	1·40	1 & 3	6·00
	3, Ring position..	1·84	1·53	2 & 3	0·70

In Table 4, the intercorrelations between the amount recalled in the sections, are all significant at better than the 2 per cent level, the smallest being between 1, (Letters) and 3, (Ring-position), and the largest between 2, (Numbers) and 3, (Ring-position). While the difference in size between these two coefficients does not quite reach the 5 per cent level of significance, this finding is consonant with the usual one on other kinds of learning task by the writer and others such as Langdon (4), that there is a tendency for test sections which are temporally nearest to be most intimately related.

Table 4

Intercorrelations between amount recalled at end of test (N = 177)

Sections	r	P
1, Letters and 2, Numbers	0·25	0·001
1, Letters and 3, Ring position	0·19	0·012
2, Numbers and 3, Ring position	0·40	<0·001

As indicated by the coefficients in Table 5, there is a significant negative relationship between the number of errors made during the learning process and the number of correct associations recalled at the end, but none between the latter and the time taken. The more efficient the learning, as reflected by the smallest number of wrong responses made in establishing correct associations, the better is the retention of what has been learned. This agrees with one of the basic principles underlying programmed instruction, namely, that the pupil learns better when making few or no mistakes.

Correlations between " Recall " and " Errors " might well have been larger, were it not for the fact that the programme consisted of three different kinds of material presented successively, which introduced variates like perseveration, negative transfer and retro-active inhibition, that would have tended to depress the correlations. A lack of relationship between the time taken to work through the sections and the amount recalled at the end, would be consonant with the view that in trial-and-error tasks of this kind rate of work may depend a lot on factors other than learning ability, and it is therefore a less pure index of learning ability than the number of wrong responses made during the task. Those advocates of self-administered programmed instruction, who maintain that the pupil learns most efficiently when he works at his own natural rate, probably have good grounds for their assumption.

Table 5

Correlations between amount recalled at end of test and Time and Errors in the learning process. (N = 177)

Sections	Time	Errors
1, Letters ..	0·01	−0·25s
2, Numbers	−0·04	−0·33s
3, Ring position..	−0·01	−0·25s
4, Total 1 + 2 + 3	−0·03	−0·38s

" s " implies significance at better than the 1 per cent level.

Validity

Relationships between performance on the test and on various flying-training criteria, are given in Table 6. All criterion scores were positively accelerated, with the exception of " Hours to solo ", where a *low* score, implying that the pupil was competent to fly alone after only a short period of dual instruction, meant good achievement. Both the " Time " and" Error " measures correlate significantly with: (a) Sport data, which comprised assessments of pupils' skill and initiative in group activities of the athletic and regimental kind; (b) Approach aids, concerned mainly with landing and target-finding by electronic techniques; and (c) Combined percentage of examination marks in both theoretical subjects and practical airmanship. The " Time " measure also correlates with Morse code, Navigation and Flying, and the " Error " measure with Radar and Hours to solo. Fast-working and/or few errors on the learning test tend to be associated with success in the criteria mentioned, the most substantial coefficients being −0·43 (Approach-aids) and −0·41 (Flying per cent) with " Time ", and 0·47 (Hours to Solo) with " Errors ".

With regard to the correlation of 0·30 between Radar and the " Recall " measure, Prof. Gourlay has pointed out that its actual significance is distinctly dubious because it is the only coefficient among fifteen that shows any significance, and it can be demonstrated by means of the binomial expansion that where fifteen correlations in all are being tested there is a 54 per cent chance of getting at least one significant (at the 5 per cent level) on the null hypothesis.

Table 6

Correlations between test measures and various Flying-Training criteria. (Decimal points omitted)

N	Flying-Training Criteria	P.A.L. Test		
		Time	Errors	Recall
57	Sport data	−36s	−30s	−02
55	Aero Dynamics	04	−04	19
57	Meteorology	−20	−18	08
55	Aviation	−09	−02	−05
55	Engines	−13	−21	−05
55	Instruments	09	−04	23
63	Radar	−11	−29s	30s
55	Approach Aids	−43s	−26s	−09
55	Morse code	−27s	−06	−02
56	Navigation	−29s	−21	−03
53	Hours to Solo	12	47s	04
53	Wings Test	−21	−21	−06
54	Ground %	−25	−17	03
55	Flying %	−41s	−22	−08
53	Combined %	−33s	−27s	00

" s " implies significance at the 5 per cent level or better.

Correlations with other tests

Table 7 shows the correlations obtained between the three measures on the learning test and various other tests. Both " Time " and " Errors " show small to moderate negative correlations with achievement on the first five, which are intellectual tests of the " pencil and paper " type, and that on one practical sensory-motor test, the Willemse-board (5), in which a long, sloping board has to be tilted left or right to control a marble so that it passes various baffles without rolling off the edge of the board except at a point where it can drop into a pocket. The largest coefficient (−0·38) is between " Time " and the Designs test where subjects had to identify the presence of a model line figure resembling a capital " Sigma " in various complex configurations of lines. As the five reference-symbols used in the learning test were also angular figures made up of straight lines it is possible that there was some overlap by a recognition factor for the kind of material. " Time " also correlates with serial-discrimination on an apparatus having eight marked keys which had to be pressed singly in response to signs appearing in a window, Dots-cancellation, " Floating Effect " co-ordination, and Variable co-ordination. The two former are mainly perceptual tests, and the two latter, sensory-motor.

There are no significant correlations between the " Recall " measure and any of the other tests, which, in addition to the fact that only one of the criteria in Table 6 showed a significant correlation with " Recall " (probably fortuitous), indicates that retention for the kind of material used in this learning test is very specific. Of the three measures, " Time " shows the widest relationship to other variables, possibly because it reflects a rather more general personal characteristic of the individual than the " power " measures.

None of the measures on the learning test shows any significant relationship to Arm-leg co-ordination specifically intended for air-pilots, Alpha-frequency on EEG, or inventories of temperament and personality tendencies.

Table 7

Correlations with other tests. (Decimal points omitted)

N	Other Tests	P.A.L. Test		
		Time	Errors	Recall
177	Mental Alertness ..	−19s	−21s	02
177	Arithmetic	−29s	−27s	08
176	Mathematics	−15s	−26s	−03
132	Gottschaldt Figures ..	−18s	−18s	16
132	Designs Test	−38s	−25s	−05
120	Dots-cancellation ..	−23s	−03	−10
177	Serial Discrim. (Time + Errors)	23s	00	−06
177	Willemse Board (Av.) ..	−23s	−23s	07
177	Floating Effect (Error − Time	22s	03	−13
177	Arm-leg Co-ord. (Error − Time)	13	03	02
177	Variable Co-ord. (Time + Errors)	25s	−06	11
118	Alpha-frequency EEG ..	14	02	−01
128	Personality Inventory ..	03	03	−11
172	Temperament Questionnaire	−01	05	−10

The means and standard deviations of subjects whose home language was respectively English or Afrikaans are given in Table 8. Although it appears that the Afrikaans group took slightly longer to work through the programme, made a few more errors and recalled a trifle less, the critical ratios of the differences between the means show that none of the differences reaches the 5 per cent level of significance.

Table 8

Respective means and standard deviations of English and Afrikaans speaking subjects

Letters + Numbers + Ring position	English N = 72		Afrikaans M = 105		C.R. of Diff. Means
	Mean	S.D.	Mean	S.D.	
TIME (secs)	614·38	116·80	646·51	166·39	1·50
ERROR	49·63	20·82	54·58	23·20	1·47
RECALL	6·53	3·28	6·19	3·19	0·68

Summary and Conclusions

1. Intercorrelations between the three sections of the paired associate learning task are higher on the " Time " than the " Error " measure, suggesting that errors provide a stronger reflection of specificity in the material than the rate of work, which is a broader individual trait.

2. Correlations between " Time " and " Errors " are large, mainly because of experimental dependence in the particular test procedure, where the making of error tended to prolong the task.

3. There were significant differences in the mean number of errors made in each section, but not in the mean time taken to complete each. Errors increased progressively from the first section to the third. Significantly more of the first section was recalled correctly than of the second or third. Possible explanations for these results are firstly, the effect of negative transfer through experimental sequence, and secondly, differences in the degree of associative affinity between stimulus items and reference symbols.

4. Sections which were temporally nearest during the learning process tended to be most intimately related in the recall task.

5. There was a significant negative relationship between the number of errors made during learning and the number of correct associations recalled at the end, but none between " Time " and " Recall ". Rate of work is probably a less pure index of learning ability than the number of wrong responses made.

6. Of several significant correlations obtained between the learning test and flying-training, the most substantial were —0·43 (Approach aids) and —0·41 (Flying per cent) with " Time ", and 0·47 (Hours to solo) with " Errors ".

7. Both " Time " and " Errors " showed small to moderate correlations with achievement on five intellectual tests and one sensory-motor test, the Willemse-board. " Time " also correlated with two other sensory-motor tests, and two perceptual tests. The " Recall " measure, which appears to be the most specific, did not correlate significantly with any other test. None of the measures on the learning test showed any relationship to Alpha-frequency on EEG or temperament and personality variates mainly of the introversion/extraversion kind.

8. On none of the measures in the learning test was there any significant difference between the mean performance of subjects whose home language was respectively, English or Afrikaans.

ACKNOWLEDGMENTS

I wish to thank Mrs. T. Retief and Miss R. Vaandrager for their great assistance in administering the test, and Prof. N. Gourlay for statistical criticism.

REFERENCES

1. DE WET, D. R. A portable hand-foot reaction test. *J. Nat. Inst. Personnel Res.*, 1960, 8, 106–116.
2. DE WET, D. R. A multiple-choice instructional device. *Psychol. Afr.*, 1964, 10, 219–235.
3. HOWIE, D. Speed and accuracy. *Aust. J. Psychol.*, 1956, 8, 111–118.
4. LANGDON, J. N. A note on the repetition of a simple motor test. *Brit. J. Psychol.*, 1931, 22, 55–61.
5. SKAWRAN, R. F. The Willemse-board as a personality test. *Psychol. Afr.*, 1962, 9, 72–85.

Mr. D. R. de Wet,
National Institute for Personnel Research,
P.O. Box 10319,
Johannesburg,
Republic of South Africa.

Manuscript received: 16th February 1966.
Revised manuscript received: 18th March 1966.

A SERIAL DISCRIMETER FOR STUDYING PERSEVERATION AND LEARNING

In some of his earlier researches on sensory-motor skills, Seashore (3) included an apparatus which he called a "Serial Discrimeter". Basically, it consisted of four finger-keys, an upright screen with a small window, and behind the screen, a weight-driven disc having four numbers repeated in random order round its edge at equal spatial intervals. Each of the keys, marked with one of the four numbers, was connected by a string to a catch that engaged pegs in the back of the disc. When the subject pressed a key, the number of which corresponded to the number on the disc exposed in the window, he released the appropriate catch and the disc was drawn by the weight to the next peg-position, exposing another number in the window. If he pressed the wrong key, the disc did not move on. His task was to react as quickly as he could to each stimulus until the whole series had been presented by a complete rotation of the disc.

Seashore (3) appears to have used the apparatus mainly to obtain a measure of simple choice reaction by direct association, but the principle also has other research applications.

The apparatus that follows, is based on Seashore's prototype; the mechanical system, whereby suitably spaced pegs on the back of a weight-driven disc engage catches connected by strings to the finger-keys, being essentially the same. However in the present design, certain additions and modifications have been introduced with the object of increasing the usefulness for the study of perserveration, learning, and problem-solving. These include, four additional keys; a pawl-and-ratchet between the reel and the disc so that the driving-weight can be wound up without having to depress all the keys; screw-rods and nuts on the keys for adjusting the strings and catches; a counter recording the number of keys depressed; adjustable stops to limit their traverse; and provision for easy interchangeability of stimulus-cards on the disc and indicator-strips over the keys.

Description

A ply-wood base-board supports two uprights joined together by a strong cross-member. (Plate 4). Two vertical pieces of flat steel bolted at the middle of the cross-member have holes near their tops to provide bearings for a transverse shaft carrying a twelve-inch ply-wood disc at one end and a winding-reel with a small handle at the other. On the inner side of the reel a spring-loaded pawl engages a ratchet pinned to the shaft, enabling the reel to rotate freely in a clock-wise direction for coiling up a length of cord that supports a lead weight via two pulleys at the top of a vertical post bolted to the wooden cross-member.

On the other side of the shaft-mounting, eight spring-loaded catches are located in a row parallel with the cross-member and each one is connected by a length of nylon fishing-line passing over a pair of small pulleys to an eyelet on a wooden key. The ends of the lines passing through the eyelets are tied to short lengths of screw-rod in metal blocks fixed behind the eyelets, and by means of nuts on the rods, slack in the lines can be taken up and the zero-setting of the catches adjusted.

On each of forty-eight equidistant radii on the back of the disc, a metal peg is fitted at the correct distance from the centre to be in alignment with one of the catches, but the order in which the catches engage pegs on successive radii as the disc moves round, is arranged to be random, to ensure that in a test series the keys do not have to be pressed successively in accordance with some easily recognized system.

The wooden keys pivot at their ends on a rod mounted near the back edge of the base-board; brass bushes, tightly press-fitted into the wood, serve also as spacers between the keys. A spring-loaded "tally-bar" pivoting on the outer ends of the rod, extends along under the keys and is linked by a light coil-spring to a small arm attached to the mechanism of a "post-office" counter fixed to the chassis. Whenever a key is depressed, the counter registers one unit.

A wooden rail with felt stops is located over the keys directly below the cross-member, to which it is attached at either end by a pair of nuts on a screw-rod. By adjusting these nuts the rail can be set higher or lower to increase or decrease the traverse of the keys between the stops here and a similar set of non-adjustable ones fixed directly to the base-board further forward. The optimum setting is that which allows an adequate but not unnecessary amount of traverse for drawing the points of the catches clear of the pegs on the disc.

At the front of the disc where the stimulus-cards are fitted (Plate 3), four short pegs are located unsymmetrically around the centre to facilitate standard positioning of the cards, which have holes punched in them to correspond with the pegs. The cards are securely held in place by a sheet-metal retainer (also perforated to fit over the pegs) and a knurled finger-nut that screws on to the projecting end of the shaft.

The front of the chassis which hinges down for fitting and removing stimulus-cards, has a small window at the top middle to expose the stimuli, and two screw-eyes engaged by the thumb-catches that lock it at the upright position. A flat metal holder for indicator-strips is attached above the finger-ends of the keys. (Plate 2).

Testing Procedures

Section 1 (Perseveration)

The card shown on the disc in Plate 3 carries the stimulus-series intended for measuring sensory-motor perseveration. There are four kinds of geometrical figures in two colours, red and black, arranged in two semi-circular groups with the word "Stop" in between. Having fitted the card to the disc, and set the latter at the starting position with the word "Stop" at segment "C" of the card at the point where it will be opposite the window, the observer raises the hinged front of the apparatus, puts the catches on, and slips the indicator-strip for this particular card into the holder over the keys. (Plates 1 and 2).

The subject is then told that red and black figures like those on the strip will appear one by one in the window, and as quickly as he can he must press the corresponding key with the forefinger of his preferred hand. Reaction on a wrong key registers on the counter, but the stimulus in the window remains stationary until the correct key is pressed. The first figure is brought into view by pressing the end key on the right.

The time taken by the subject to complete the first trial, and the number of wrong responses he has made, are noted. Errors are indicated on the counter as the number recorded in excess of twenty-four, which is the minimum tally if all the reactions (including that on the right-end key for starting) have been done correctly.

In the second trial, when a black figure appears in the window, the subject must press the red key marked with the same shape, and when a red one appears, its black equivalent. But when a figure in the window has a dot under it, then he must press the key marked with the same shape and colour, as he did in the first series. Time taken and number of wrong responses made is noted here also. The third trial is a repetition of the first (Direct Association), and the fourth, of the second (Colour Change).

Section 2 (Learning)

An indicator-strip with the four outer keys marked by numbers but the four middle ones left blank (Plate 2) is placed in the holder, and a card having numbers corresponding to these and also four kinds of black geometrical figures round its circumference (Plate 2) is fitted to the disc.

When a number appears in the window, the subject simply presses the key marked with the corresponding number. When a figure appears, he starts by pressing any one of the four unmarked keys. If the stimulus moves on, he has selected the correct key for that particular figure. If the stimulus does not move on, he has not found the correct key and must try another. Once he knows that a certain key is correct for say, the triangle, then he must always react on it whenever the triagle appears. Similarly, the key which causes the square to move on, is always correct for the square, and so forth. He must try to master the combination as quickly as he can, making the smallest number of errors. Three identical trials are given in succession, the time and number of errors for each being noted.

Results and Discussion

Both the "Perserveration" and "Learning" sections of the test were administered experimentally to a group of 161 air-pilot candidates who were all physically-fit matriculants with normal vision.

Table 1 gives the mean Time and Error scores for each of the four trials in Section I (Perseveration). As could be expected, more time is taken and more error made in Trials 2 and 4 requiring the additional cognitive process of sometimes having to reverse the normal colour association between response and stimulus, and sometimes, not.

The intercorrelations between the four trials in Section I (Perseveration) are shown in Table 2. In general, the coefficients on the Time measure tend to be longer (mean r, 0.61) than those on the Error measure (mean r, 0.37), although between the first and second trial they are the same, namely 0.58. The Error measure shows less internal consistency, because the distribution of errors in this section of the test was rather sparse, and unsystematic influences had therefore more scope in the Error measure. The longest correlation on the Error measure is between Trials 1 and 2 where the novelty of the test situation would have tended to favour the production of errors. On the Time measure, the longest coefficients are between identical trials: 1 and 3 Direct Association, 0.65; 2 and 4 Colour Change, 0.70. On the Error measure, one of the longest is between identical trials: 1 and 3 Direct Association, 0.46. The effect of perseveration in the broad sense of lag or inertia in establishing the correct association between stimulus and response, while most marked in the "Colour Change" trials, is probably also present to some extent in the cognitively simpler activities in the "Direct Association" trials. Furthermore, when a "Direct Association" trial such as the third, follows immediately after a "Colour Change" trial, then the latter would exert a moderate hindrance effect on the former.

The correlation between Total Error and Total Time in Section I (not shown in the table) is 0.38. Nothing of behavioural interest can be inferred from this because it is mainly attributable to some experimental dependence between the measures. Each stimulus remained in position until the subject reacted on the appropriate key, and mistakes therefore, tended to retard the rate of work.

Table 3 gives the mean Error and Time scores for each of the three trials in Section 2 (Learning). The effect of learning is apparent in the progressive decrease in Error and Time on successive trials.

In Table 4, the intercorrelations between trials are generally higher than in Section I (Table 2) because all the trials, (with particular reference to those adjacent in the test), were identical, and moreover, twice as many stimuli were presented in each, as in Section 1. The Error measure shows markedly greater internal consistency than that in Section I, for, in addition to all the trials being identical and longer, the greater complexity of the learning task induced many more errors.

The correlation between Total Error and Total Time in Section 2, Learning, (not shown in table) is 0.42. Here, also, the result is largely due to experimental dependence between the measures, as wrong responses caused delays.

The following correlations were obtained between the total scores on Section I (Perseveration) and Section 2 (Learning): Errors, 0.22; Time, 0.55. (N = 161). Although, in Section 2 (Learning) there were many more errors than in Section 1 (Perseveration), errors as a whole still constitute a comparatively sparsely distributed measure, and this would likewise have played a part in depressing the "Error" correlation here. There is also the possibility that "Error" being a power measure, is more dependent on specific aspects of the task in either section, than the Time measure, which shows a large correlation because rate of work is a broader personal characteristic of the individual.

Of the 161 air-pilot candidates who did the Discrimeter test, forty were selected on the basis of other test procedures (excluding the Discrimeter) to undergo flying training. Later, thirty-two of this group succeeded in passing the flying course and eight failed through poor flying ability.

Biserial correlations obtained between the Pass/Fail criterion and total scores in the Perseveration and Learning sections of the Discrimeter are given in Table 5. In neither section does the Time measure show any significant correlation, but the Error measure shows significant correlations in both: Perseveration, 0.46 (P = 0.01); Learning, 0.38 (P = 0.05).

In this study, the hypothesis that success in learning to pilot an aircraft is associated with low perseveration and good learning ability in a perceptual-motor situation, is supported only by the results on the Error measure. However, in a subsequent investigation on paired associate learning with other apparatus (2), results indicated that both Time and Error measures in the learning tasks were related to certain branches of flying training.

Relationships between the Perseveration and Learning sections on the Discrimeter and other practical tests are shown in Table 6. Error in Section I (Perseveration) correlates with nothing else except Arm-leg Co-ordination (with distraction) where the coefficient (0.17) is probably just a chance result. In Section 2 (Learning) the small correlations between Error and Two-hand Co-ordination (Moede Type, Time) and Hand-foot Reaction (Learning Time) are likewise, barely significant. Time, in Section 1 (Perseveration), correlates with Two-hand Co-ordination (Moede Type, Time), 0.37; Arm-leg Co-ordination (with Distraction) 0.35; and Hand-foot Reaction (Reaction Time), 0.52. Time, in Section 2 (Learning) correlates with Two-hand Co-ordination (Moede Type, Time), 0.28; Arm-leg Co-ordination (with Distraction), 0.21; Manual Steadiness (Error), 0.17; Hand-foot Reaction (Reaction Time), 0.47; and Handlebars (Error), 0.21.

On both the Perseveration and Learning sections, Time shows a wider overlap with the other tests than Error, which, in view of the intercorrelations between the Discrimeter test trials in Tables 2 and 4, where those for Time tend to be the larger, lends support to the theory that, in general, correlations on Error are smaller because (a) The distribution of errors is comparatively sparse., and (b) Error is a more intrinsic reflection of task specificity, whereas rate of work is a broader personal characteristic of the individual. Similar results were obtained in the test of paired-associate learning (2).

There is a notable lack of relationship between the present Discrimeter measures and Flicker-fusion threshold (1). It is therefore to be concluded that perseveration or secondary-function on the Discrimeter, reflected by a tendency to work slowly and make wrong responses (particularly under conditions requiring the reversal of normal associations) is not significantly related to the perseveration or secondary function in Flicker-fusion, manifested in the perception of fusion of discrete visual stimuli presented at comparatively low frequencies.

Table 7 gives the correlations obtained on two later annual groups of candidates, between a combined speed and accuracy score on Section I (Perseveration) only, and achievement on various other tests. The Perseveration score consisted of total Time for the four trials, plus total Error multiplied by ten. In Group 1, there are small correlations with three printed tests: Mental Alertness, Technical and Scientific Information, and Repeated Letters; and with three apparatus tests: Span of Attention, Willemse Board (co-ordination) and Paired Associate Learning (Speed). In Group 2, the composition of which, as regards age, sex, educational level, intelligence, and cultural background, was similar to that of Group 1, there are no relationships between Perseveration scores and any of the above variates. (Paired associate Learning was not administered). Here there are small correlations with only Arithmetic and Gottschaldt Figures.

These results demonstrate, that very small relationships, although statistically significant, should be regarded with particular caution, because they can fluctuate markedly, even between samples that are very similar in composition.

Correlations between the combined speed and accuracy score in Section I (Perseveration) and various branches of a course in flying training, are shown in Table 8. The only ones significant at the 5% level or better are Aviation (0.25), Athletic Sports, Games, etc. (0.36) and General Service Knowledge (0.38). Low perseveration, as reflected by fast working and few mistakes on the test, would appear to be associated slightly with success in these branches. The relationships of 0.23 with Night Flying and overall achievement in the Ground or theoretical subjects, are significant at the 10% level. On this sample there is no correlation with the Pass/Fail criterion. It is possible that the present combined score of Time and Error is unsatisfactory as it gives undue weight to the Time measure, which in the first validation study (Table 5) showed no useful relationship to flying success.

The mean combined Perseveration scores, and also the mean scores in some other tests, of subjects whose home language was respectively, English or Afrikaans, are given in Table 9.

In the three intellectual tests, namely, Mental Alertness, Technical and Scientific Information and Gottschaldt Figures, the English group did better; particularly in Technical and Scientific Information, where English-speaking subjects would have had the advantage of wider previous exposure to technical and scientific literature in their own language. In Span of Attention and Paired Associate Learning there are no significant differences between the groups. In Repeated Letters and Dots Cancellation, which are mainly simple perceptual tasks requiring sustained attention, the Afrikaans group was superior. A similar result has been obtained by Dr. Reuning of this Institute with another test of sustained attention, the Pauli. In the Perseveration section on the Discrimeter, the English group did slightly better, the difference being significant at the 5% level, but in Flicker-fusion this group was actually the more secondary, which is a further indication that the aspects of secondary function measured by these two tests are largely independent.

Summary and Conclusions

1. An improved Discrimination apparatus, inspired by an earlier design of Seashore's, but intended mainly for studying perceptual-motor perseveration and paired-associate learning, is described.
2. In Section I (Perseveration), subjects tended to take longer and make more wrong responses when they had to deal with changes in the normal colour association between stimulus and response.
3. The Time measure showed greater internal consistency than the Error measure, which, being much more sparsely distributed, was more subject to unsystematic influences.
4. In Section 2 (Learning) there was a progressive decrease in Time and Error on successive trials. Inter-trial correlations were generally higher than in Section I because all trials were identical and errors occurred more frequently.
5. The correlations between the Perseveration and Learning sections were: Errors, 0.22; Time, 0.55.

6. Only the Error measure in either section showed a significant biserial correlation with the Pass/Fail flying criterion: Perseveration, 0.46 (P = 0.01); Learning, 0.38 (P = 0.05).
7. On both the Perseveration and Learning sections, Time showed a wider overlap with other practical tests than Error which, in addition to its comparatively sparse distribution and hence greater susceptibility to unsystematic influences, is probably a more intrinsic reflection of task specificity. Rate of work is a broader personal characteristic of the individual.
8. No significant relationship was found between the present measures of Perseveration on the Discrimeter and Flicker-fusion threhold. It would seem therefore, that the aspects of secondary-function measured in the respective tests are not closely connected.
9. Small correlations between a combined speed and accuracy score of Perseveration and achievement on certain other tests (mainly intellectual and perceptual), were not consistent in two similar groups of subjects.
10. The combined Perseveration score correlated significantly with the following branches of a course in flying training: Aviation, 0.25; Athletic Sport, Games, etc., 0.36; General Service Knowledge, 0.38.
11. English-speaking subjects fared better in three intellectual tests, and also to a slight but significant extent, in the Perseveration test, although they rated as more secondary in Flicker-fusion. Afrikaans-speaking subjects scored more in Repeated Letters and Dots Cancellation. In Span of Attention

and Paired Associate Learning differences between the groups were not significant.

12. In the present study, three perseveration measures were investigated in relation to other variates: Total Time for both the "Direct Association" and "Colour Change" activities, Total Error for both, and a combination of Time and Error. This was done on the rationale that perseveration in its wider sense as a form of temperamental inertia could be expected to influence performance in both activities. However, perseveration in its restricted sense of hindrance effect exercised by previously established associations over the acquisition of new ones, is more marked in the "Colour-Change" activity, as shown by the means in Table I. It is hoped, later, to do another study using as perseveration measures the differences between the scores in the "Direct Association" and "Colour Change" activities, and scores in the latter, only.

13. Learning tasks can be complicated by marking fewer keys on the indicator-strip, and will be most difficult when there is no indicator-strip at all, the subject being required to learn the association between each of the eight keys and a stimulus. The apparatus can also be used for presenting a "problem-solving" situation in which success depends on recognizing some systematic mathematical or other relationship between numbers, symbols, etc., appearing in the window and other configurations identifying the keys.

The stimuli on the cards can be in the form of numbers, letters, figures, signs, colours, etc., either separate or in combination, provided the overall size of each is within the limits of the window on the apparatus.

ACKNOWLEDGMENTS

I am grateful to Miss M.E. Jones and other staff members for their assistance in test administration.

REFERENCES

1. DE WET, D.R. A compact flicker-fusion machine and its application to air-pilot candidates. Psychol. Afr., 1962, 9, 100 - 118.

2. DE WET, D.R. Paired-associate learning through a practical medium. Psychol. Afr., 1966,

3. SEASHORE, R.H. The Stanford motor skills unit. Psychol. Monogr., 1928, 39, No. 2, 51-66 (Whole No. 178)

TABLE 1

Mean Time and Error Scores on each of the four trials in Section 1 (Perseveration)

N=161	1. Direct Association	2. Colour Change	3. Direct Association	4. Colour Change
ERRORS				
Mean	1.69	2.87	1.78	2.25
S.D.	1.84	2.50	1.34	1.39
TIME (Secs.)				
Mean	31.79	43.35	28.73	38.23
S.D.	4.40	6.86	3.58	5.59

TABLE 2

Intercorrelations between test trials in Section 1 (Perseveration)

Trials N = 161	Errors	Time
1 (Direct Assoc.) and 2 (Colour Change)	0.58	0.58
1 (Direct Assoc.) and 3 (Direct Assoc.)	0.46	0.65
1 (Direct Assoc.) and 4 (Colour Change)	0.37	0.51
2 (Colour Change) and 3 (Direct Assoc.)	0.22	0.60
2 (Colour Change) and 4 (Colour Change)	0.36	0.70
3 (Direct Assoc.) and 4 (Colour Change)	0.20	0.60
Mean r	0.37	0.61

TABLE 3

Mean Time and Error scores on the three trials in Section 2 (Learning)

N = 161	1	2	3
ERRORS			
Mean	24.75	15.23	10.60
S.D.	10.92	10.94	9.81
TIME (Secs.)			
Mean	88.35	70.90	64.17
S.D.	15.66	12.95	12.06

TABLE 4

Intercorrelations between the three trials in Section 2 (Learning)

Trials N = 161	Error	Time
1 and 2	0.60	0.69
1 and 3	0.54	0.60
2 and 3	0.79	0.83
Mean r	0.64	0.71

TABLE 5

Biserial correlations with Pass/Fail criterion

TEST SECTION	r bis	P
1. PERSEVERATION		
Trials 1 + 2 + 3 + 4 (Time)	0.15	0.45
Trials 1 + 2 + 3 + 4 (Errors)	0.46	0.01
2. LEARNING		
Trials 1 + 2 + 3 (Time)	0.18	0.40
Trials 1 + 2 + 3 (Errors)	0.38	0.05
N = 40 (Passes 32, Failures 8)		

TABLE 6

Correlations between Perseveration and Learning Sections
and other tests
(Decimal points omitted)

Other Tests N = 161	Sect. 1 Perseveration		Sect. 2 Learning	
	Error	Time	Error	Time
Flicker Fusion	01	13	-08	-01
Two-hand Co-ord.(Moede Type) Error	13	08	08	08
Two-hand Co-ord. (Moede Type) Time	13	37s	17s	28s
Arm-Leg Co-ord. (Section 1)	-01	12	06	07
Arm-Leg Co-ord. (With Distraction)	17s	35s	-03	21s
Manual Steadiness (Time)	07	05	00	-01
Manual Steadiness (Error)	-09	11	15	17s
Hand-foot Reaction (Learn.Time)	02	14	18s	12
Hand-foot Reaction (Reaction Time)	03	52s	-07	47s
Hand-foot Reaction (Error)	07	-05	12	-08
Handlebars (Time)	05	-01	-06	-07
Handlebars (Error)	03	14	08	21s

"s" implies significance at the 5% level or better

TABLE 7

Correlations between a combined speed and accuracy score in Section 1 (Perseveration) and other tests.
(Decimal points omitted)

OTHER TESTS	Group 1		Group 2	
	N	r	N	r
Mental Alertness	177	19s	124	09
Arithmetic	177	08	124	19s
Tech. and Scientific Information	177	18s	124	01
Tech. and Scientific Read Comp.	177	04	124	05
Science (Matric)	176	03	124	12
Mathematics (Matric)	176	04	125	13
Gottschaldt Figures	132	-01	122	21s
Designs Test	132	10	122	07
Dots Cancellation	120	08	122	06
Repeated Letters	132	18s	122	08
Span of Attention	177	24s	128	02
Personality Inventory	128	03	124	02
Temperament Questionnaire	172	-05	124	-13
EEG Alpha Frequency	118	-07	128	10
Floating Effect (co-ord.)	177	-04	128	-05
Variable Co-ordination	177	02	128	12
Willemse Board (Average)	177	15s	128	-02
Paired Assoc. Learn. (Speed)	177	23s	-	-

TABLE 8

Correlations between a combined speed and accuracy score in Section 1 (Perseveration) and various flying-training criteria.

Flying Training Criteria	N	r
Aero Dynamics	49	-0.05
Meteorology	49	0.18
Aviation	58	0.25s
Instruments	58	-0.02
Radar	68	0.18
Approach Aids	58	0.07
Morse	51	-0.01
Navigation	50	0.15
Night Flying	50	0.23
Wings Test	49	-0.04
Ground %	49	0.23
Flying %	49	0.09
Combined %	49	0.17
Total Pass/Fail	69	0.08
Athletic Sport, Games etc.	49	0.36s
General Service Knowledge	51	0.38s
Character	50	0.20

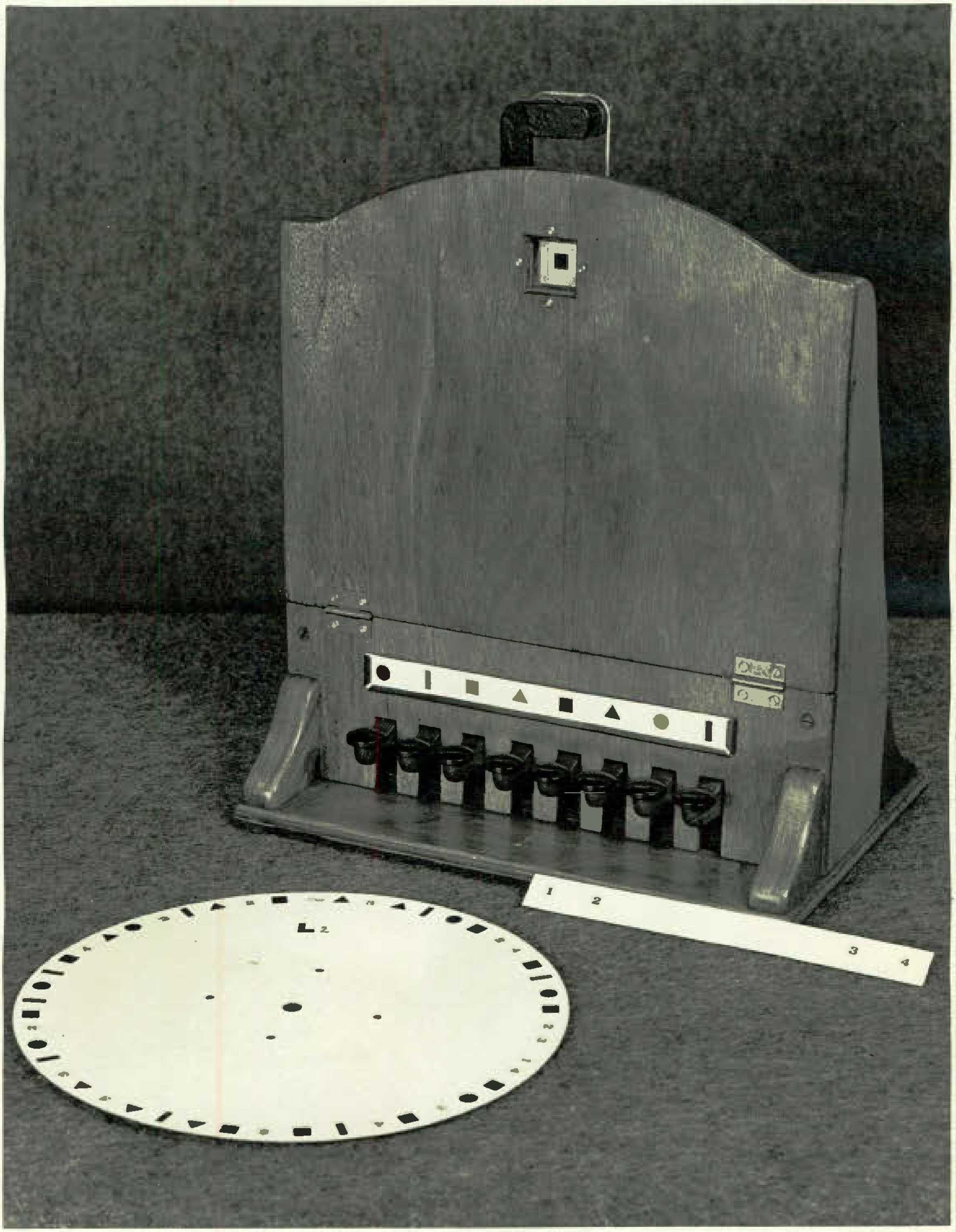
TABLE 9

Mean scores of English and Afrikaans speaking subjects

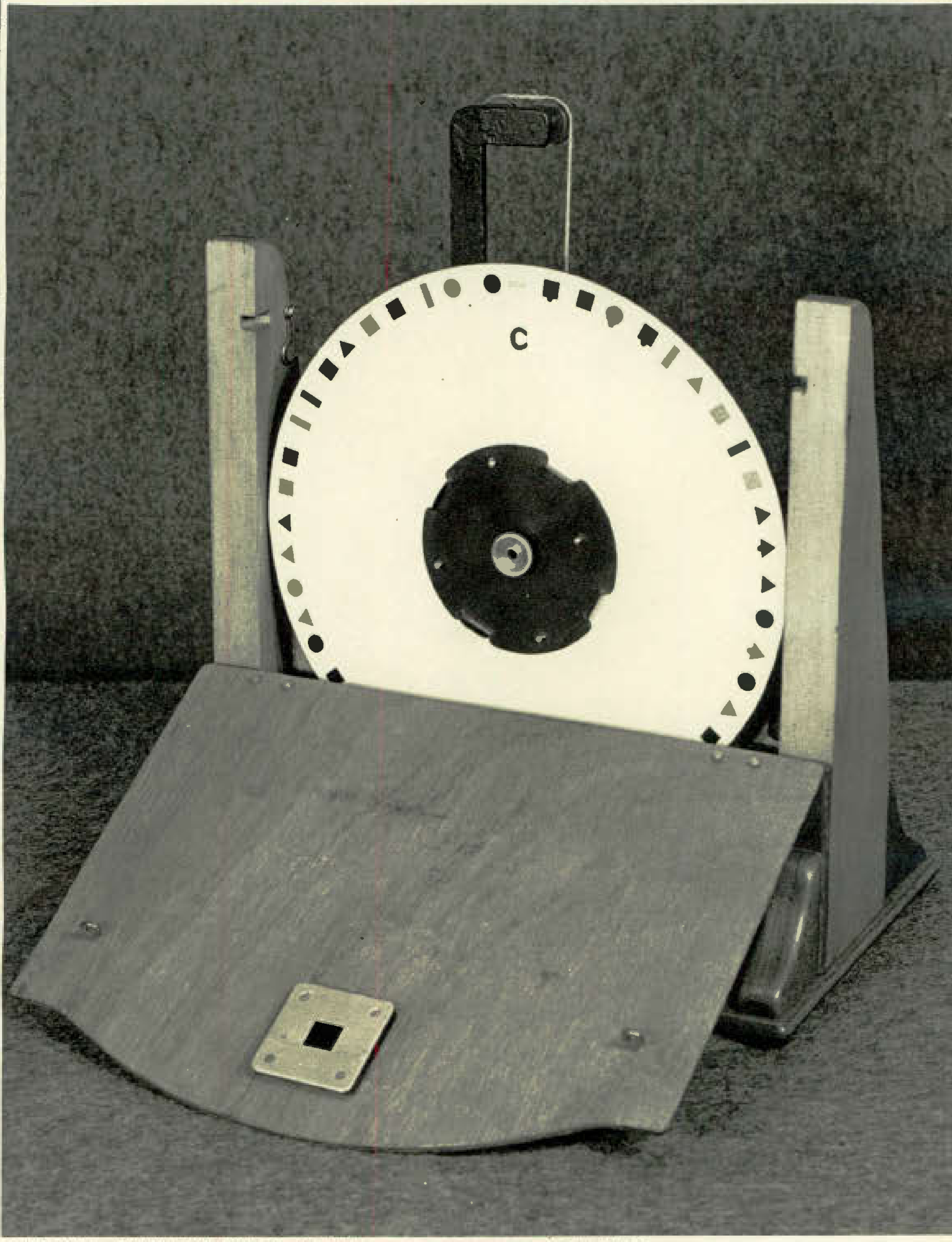
TESTS	N	English Mean	English S.D.	N	Afrikaans Mean	Afrikaans S.D.	C.R. of Diff.M's
Mental Alertness	72	27.76	3.98	105	24.49	5.27	4.68
Tech. and Scient. Inform.	72	37.49	4.46	105	29.99	5.44	9.98
Gottschaldt Figures	59	26.86	8.13	73	21.58	8.54	3.61
Span of Attention	72	43.35	5.99	105	42.76	6.43	0.62
Flicker Fusion	72	259.49	37.64	105	242.67	44.37	2.70
Paired Assoc.Learn (Time)	72	614.38	116.80	105	649.51	166.39	1.50
Repeated Letters	59	33.83	8.40	73	41.08	9.04	4.73
Dots Cancellation	55	225.67	50.93	65	244.32	43.04	2.13
Discrimeter (Persev.Sect.)	72	151.43	20.22	105	158.85	29.40	1.98

EIGHT-KEY SERIAL DISCRIMETER. PLATE 1.

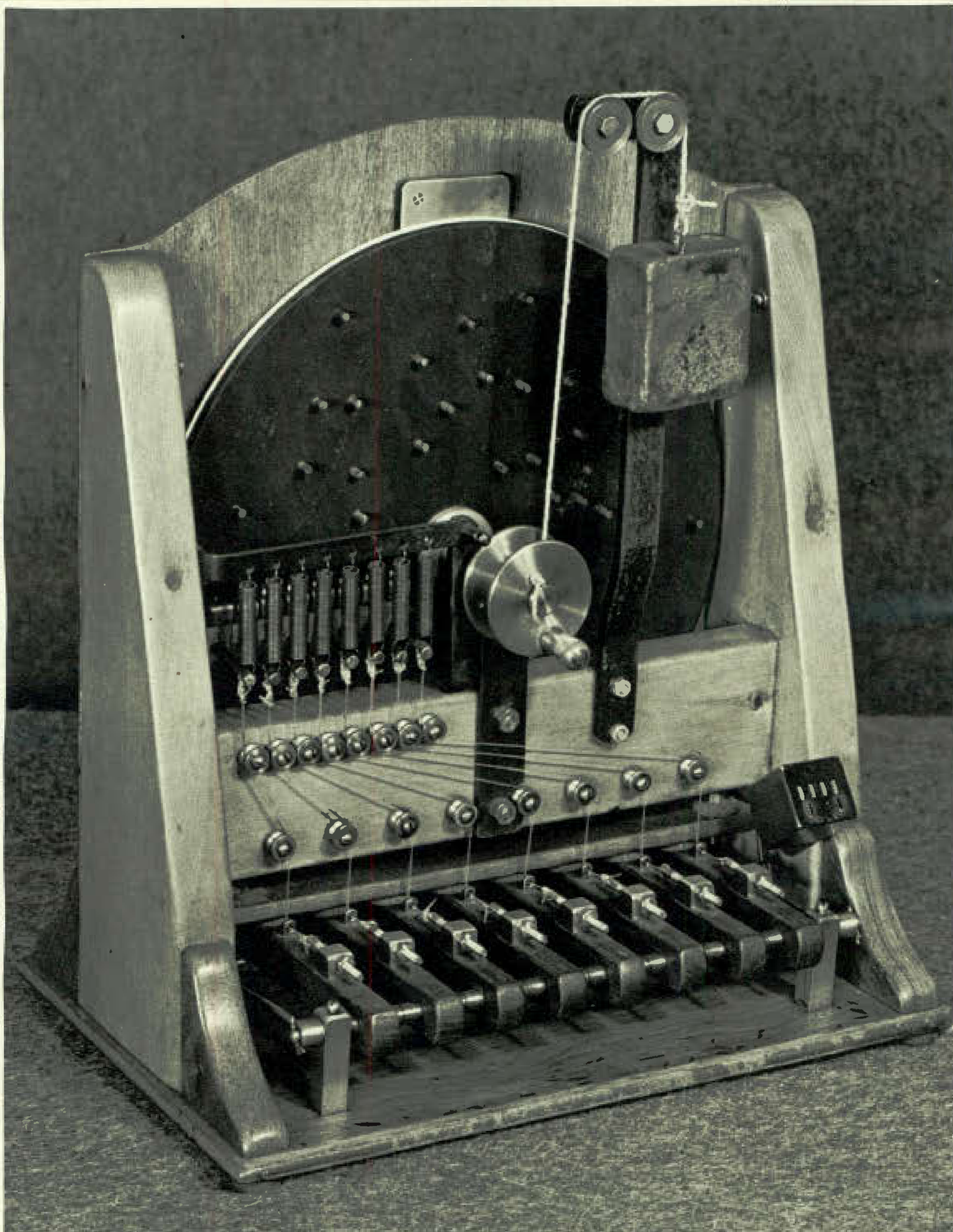
EIGHT-KEY SERIAL DISCRIMETER. PLATE 2.



EIGHT-KEY SERIAL DISCRIMETER. PLATE 3.



EIGHT-KEY SERIAL DISCRIMETER. PLATE 4.



SECTION VI

Summary

1. Technical and research accounts are given of two basic devices that provide useful practical media for investigating learning processes and other abstract functions, and also for administering various selection tests of a cognitive kind. Both have the advantage of utilizing gravity instead of electricity in their operation.

2. By means of the Multiple-choice Instructional Device, a test of paired-associate learning consisting of three distinct sections was applied to air-pilot candidates, on the assumption that learning by trial-and-error plays a significant rôle in practical flying training.

3. It was not possible to assess reliability as the test was given once only and certain differences in the content of the respective sections precluded these from being treated as similar "trials".

4. Validation results indicate that the test has potential usefulness for predicting success in certain branches of flying training, particularly, the more practical. The most significant correlations obtained were -0.43 (Approach aids) and -0.41 (Flying per cent) with "Time"; and 0.47 (Hours to solo) with "Errors".

5. Both "Time" and "Errors" showed small to moderate correlations with achievement on five intellectual tests and one sensory-motor test. "Time" also correlated with

two other sensory-motor tests, and two perceptual tests. The "Recall" measure was very specific and showed hardly any relationship to other tests or the criteria.

6. Other results from this investigation suggest that : Errors provide a stronger reflection of specificity in the material than the rate of work, which is a broader individual trait, and also a less pure index of learning ability then the number of wrong responses made. Learning was influenced by both negative transfer through experimental sequence, and differences in the degree of associative affinity between stimulus items and response references. Material that was temporally nearest during the learning process tended to be most intimately associated in the recall task.

7. Using the Serial Discrimeter, based on an earlier apparatus of R.H. Seashore's, tests of perceptual-motor perseveration and paired-associate learning were applied to air-pilot candidates. Here also, Time measures showed better inter-trial reliability than Error measures, which, being more sparsely distributed, were more subject to unsystematic influences. The general effect of Perseveration emerged as a tendency for subjects to take longer and make more mistakes when they had to deal with a substitution task involving changes in the normal colour association between stimuli and responses. Error-proneness in this activity correlated 0.46 ($p = 0.01$) with Suspension in flying training. In the Learning section Error correlated 0.38 ($p = 0.05$) with Suspension. The correlations between Perseveration and Learning were : Errors, 0.22; Time, 0.55. A perseveration score in which both

SECTION VII

PRACTICAL TESTS IN PERSONNEL SELECTION

PRACTICAL TESTS IN PERSONNEL SELECTION

Introduction

From earliest infancy we start an acquaintance with material things at a sensory-motor level, and our first ideas are intimately concerned with these physical models. To quote meredith (47), "We handle them, squeeze them, bite them, get squeaks out of them, throw them round and use them to establish relations with people". Later we associate verbal symbols with material things and learn the concept of number by manipulating concrete units, and still later we may be largely dependent on some form of concrete demonstration model in order to understand an involved principle of astronomy, chemistry, physics or mathematics. The philosopher, to clarify his exposition of complex tenets, makes frequent reference to his "table" - that good, homely piece of furniture, standing out like a solid platform of refuge to the poor student adrift on the desolate ocean of metaphysical abstraction. Even a genius like Newton was once indebted to a simple physical object in the form of an apple for providing the stimulus to some tremendous mental activity. Throughout life, for most people, concrete things remain fundamental to abstractions.

An inspection of an extensive list of modern occupations, such as that compiled by Beckman (8) who based his particular classification on the rating of occupational status according to the amount of general prestige attached to the workers in various spheres, discloses the interesting fact that, to a varying extent, most occupations are concerned with the direct or indirect manipulation of concrete things. Those restricted mainly to the management of ideas and symbols appear to be in the minority. Not only in the multifarious ranks of the unskilled, semi-skilled, and skilled manual occupations, but also at the sub-professional, minor-supervisory, and even professional levels, a large proportion of humanity makes a living by handling material things or regulating its thinking in relation to the physical arrangement of material things. Labourers, porters, messengers and cleaners have to work with both their brains and their hands, and so do engineers, artists, sculptors, surgeons, and executive musicians. The "practical" and "spatial-mechanical" factors identified by Cox (17), Vernon (72, 73) and others play a significant rôle in many diverse strata of the occupational hierarchy, and there is certainly some justification for the view, expressed in a rather startling and provocative form by Ramanathan (56), that "... the whole edifice of human civilization has been built up through the instrumentality of the human hand". Yet, underlying this instrumentality, there is, of course, the guidance of the human mind.

Personnel psychologists make a broad, working distinction between two kinds of selection test : (1) General-purpose, and (2) Specific. The first comprises tests intended to measure what appear to be the more fundamental attributes, like general intelligence; abstract reasoning; spatial and verbal ability and other major group factors, that, to a varying degree, play a part in many different occupations; and the second, tests measuring attributes that are largely unique to a particular occupation or a group of similar occupations. Either of the above test categories can be sub-divided into : (a) Those of the pencil-and-paper (or pictorial) kind, and (b) those which are administered through the medium of more concrete apparatus.

The following are typical examples of this broad classification : (1a) Mental Alertness, Otis Group Intelligence Scale, Raven's Matrices, Arithmetic, and Reading Comprehension.
(1b) Manipulative puzzles, Assembly tasks with the emphasis on insight, Form-boards, Cube-construction, and Object-sorting.
(2a) Printed tests of electrical circuitry and workshop knowledge, and interest inventories for specific occupations. (2b) Air-pilots' Arm-leg Co-ordination and Floating-effect Tests (Air-pilots), Locus-estimation Test (Air-traffic controllers), and the Code-signal and Dynamic Coincident Reaction Tests (Mine winch-operators).

For reasons of economy and convenience, the use of printed tests is favoured, wherever possible. They can be more easily reproduced than apparatus tests, are better suited to group administration, require no specialized skill to administer, and are not subject to physical variation between copies that can affect the uniformity of the task to be performed. It is also expedient to make full use of the general-purpose sort, as these are usually well-standardized basic tools always ready for immediate application. The specific test, as Biesheuvel (13) has observed, "imposes greater burdens on research resources", and should therefore only be administered when the occupation for which selection is to be done has a highly important specific element that cannot be adequately covered by general-purpose tests. This applies particularly to specific apparatus tests that require individual administration.

When selection is done on large population samples, to ensure that no costly detailed testing is wasted on candidates who are generally unsuitable, it is economical practice to pre-screen all candidates initially on general-purpose tests, and administer specific individual tests only to those whose general standard surpasses a certain minimum. This is known as two-stage sequential selection (Cronbach and Gleser, 20) and was the procedure usually employed in selecting air-pilot candidates. Pre-screening of all applicants was done on a short general-purpose battery of written intellectual tests because a certain amount of intellectual ability is essential for the pupil-pilot, particularly in mastering the theoretical subjects. Those who passed this initial screen were then tested for other requirements in the air-pilot's vocation, such as qualities of leadership, good muscular co-ordination, ability to learn in practical situations, and the right kind of temperament.

The principal consideration determining the level of rejection in the pre-screening test is the degree of relationship (validity) that the test is known to have with the occupational requirements. If selection is being done for candidature in one of the learned professions where success usually correlates highly with intelligence, then the level of rejection on a screening test of intellectual ability would fittingly be higher than for air-pilots whose vocation, in general, is not primarily intellectual, although achievement in certain of the theoretical subjects may correlate quite highly with intelligence.

Another consideration affecting the choice of a rejection level in the screening test is the relative proportion of the number of candidates who must be selected and the total number available from which to select. This is known as the Selection Ratio (S.R.). The selector seeks to locate the rejection level at a sufficiently high position which will exclude those candidates who are likely to fail in the occupation, but not so high that there is a risk of rejecting too many who might well succeed because they are adequately endowed with the other necessary qualities or attributes to be measured in the battery of more specific tests that will follow.

A pre-screening technique that can effect appreciable economy in test administration has been developed by Arbous and Sichel (2). Using a similar procedure, Gotsman (32) showed that the N.I.P.R. tests for classifying African industrial workers could be decreased in number without markedly reducing their predictive usefulness. Other "sequential strategies" are described by Cronbach and Glaser (20; p. 69).

When the sample for processing is small, the time and labour to be saved by pre-screening with the general-purpose tests is of lesser importance, and it is preferable to administer both general-purpose and specific individual tests to all candidates and assess occupational suitability in terms of a composite score in which the respective tests carry appropriate weights. This procedure, while less economical, is rather more efficient because it eliminates the risk of rejecting, through pre-screening, some candidates who might perform well on the subsequent specific tests that measure other attributes connected with success in the occupation.

However, as Bechtoldt (6; p.1238) remarks, it is seldom in personnel selection that ready-made techniques can be employed.

"The personnel psychologist must meet most situations armed only with a method of approach. He must first study the problem in terms of the purpose of the selection, then formulate working hypotheses, devise the necessary procedures, try them out experimentally, check on the validity of his hypotheses, and finally revise his procedures in accordance with his findings."

Because apparatus tests are less economical and convenient instruments for selection purposes than printed tests, attempts have been made to evolve printed substitutes. Fleishman and Hempel (27) who are notable in this field, have achieved considerable success; however, they conclude that although some of the factors measured by psychomotor tests may also be measured by printed tests, there are others that are not sampled in printed tests.

In studying the performance of engineering apprentices, Ross (61), McMahon (45) and Montgomery (50) found that general purpose tests of Arithmetic, Mathematical Achievement, Mechanical Comprehension, and Verbal Abstraction gave respective multiple correlations of 0.69 and 0.85 with a scaled supervisors' rating and a technical theory qualification, but neither the tests nor the above training assessments related significantly to ratings of candidates' subsequent performance in the practical factory jobs. Tromp (71) administered similar general purpose tests to Air-force artisan apprentices and found that there was little connection between the supervisors' ratings of practical proficiency and either the test scores or grades in theory examinations. A printed test of Mechanical Comprehension widely employed in France for the selection of engineering apprentices was investigated by Patin and Nodiot (55), who obtained correlations of

0.25 and 0.45 with final examination grades including workshop practicals, but a coefficient of only 0.22 with practical proficiency.

Scope and Limitation

While it seems logical that practical abilities can be sampled most thoroughly by means of practical test situations, usually requiring "apparatus", such tests are generally only worthwhile if they are constructionally simple, physically compact, inexpensive to reproduce, and easy to administer and score. The use of more elaborate apparatus should be reserved for the selection of high-level specialists and personnel on whose competence and dependability the safety of life and valuable property depends.

It is noteworthy that most of the apparatus tests which have survived the proof of time and continue to be regarded as useful tools for selection, fulfil the above requirements of "simplicity". Examples of these are: Koh's Block-design and the Cube-construction test, providing measures of 'g' through a physical manipulative medium, and suitable for use on candidates of widely different cultural background; the Wiggly-block test of the ability to visualise three-dimensional structure needed in architecture, engineering, pattern-making, sculpture, etc. (Keane and O'Connor, 42; Remmers and Schell, 57); O'Connor's finger-dexterity test, a useful predictor of aptitude for manipulating small components (Hines and O'Connor, 36); and the Minnesota mechanical and spatial-relations tests (Paterson et al., 54). Among such "simple"devices could also be included the type of Two-hand co-ordination test introduced by Moede (49), which, although comparatively more expensive to construct, provides one of the most useful general-purpose tests of muscular co-ordination ever devised. In principle, it reproduces the movements of a compound slide on a lathe, and was originally intended as a specific test for the selection of candidates most likely to succeed as lathe operators. However, it has also shown validity for many other occupations requiring fine, co-ordinated movements, like draughtsmanship, pattern-making, fitting, tool-making, and the operation of various machines. Although the controls of the Moede-type apparatus bear no direct resemblance to those in an aircraft, Biesheuvel (13) found that performance on this test related significantly to success in flying training.

Apparatus tests ordinarily used for selection are intended to measure either some form of sensory-motor skill, or the ability to bring material objects into some orderly relationship with one another, which may be static, as in the "formboard" type of test, or dynamic, as in a mechanical assembly that must be capable of transmitting movement. In the spatial-mechanical group, which includes tests of the "manipulative puzzle" sort intended to provide measures of 'g' through the medium of a concrete setting, while there is also a varying amount of manipulative skill involved, it is usually of lesser importance than the correct idea in mind of how things are related, spatially and functionally. Apparatus tests of perceptual ability are less commonly used, being confined to selection for those occupations where a good visual, auditory, or tactile sense is regarded as particularly necessary.

The greater basic inconvenience and expense of apparatus tests as compared with printed tests is aggravated by the fact that practical abilities, especially those requiring sensory-motor skill, tend to be very specific. Modelling a piece of sculpture, working a lathe, wrapping bon-bons, flying an aircraft, wielding a tennis-racquet, playing the violin, doing a surgical operation - all involve "manual skill", but it is largely unique to each occupation. The limitation of specificity, added to their other disadvantages, restricts the use of practical tests in occupational selection, and it is hardly surprising, that ordinarily, in the tool-kit of the personnel selector, pencil-and-paper tests far outnumber apparatus tests. One printed general-purpose test of cognitive ability can cover a wide variety of different occupations, whereas the usefulness of a test of sensory-motor skill is usually confined to a very narrow range. It is seldom that such a test is found to have a degree of versatility comparable to that shown by the Moede-type Two-hand co-ordination machine.

The following are a few examples of numerous investigations that have demonstrated the specific nature of practical skills : Seashore (64) applied the Stanford Motor-skills tests to 50 prospective winding-machine operators, and obtained a correlation of only 0.14 ($\pm$0.10) between the battery as a whole and success in working the winding machines. Even a very low score on any one of the tests was no indication that the candidate would not succeed in the job. Walker and Adams (75), who administered the same tests to students, found no significant correlation between test performance and success in learning to use a typewriter. The writer's simple skill tests applied to African personnel at a porcelain factory showed no relationship to subsequent ratings of proficiency on the job. Biesheuvel (13) has reported that in testing groups of artisan apprentices for radio and engineering trades in the South African Air Force, it was found that "the psychomotor tests do not contribute anything to the prediction of either the written or the practical examination"; and in the case of another project, conducted on groups of apprentice lift-mechanics, that the omission of the apparatus tests "only caused a shrinkage of multiple R from 0.80 to 0.75." The addition of costly individual sensory-motor tests to a relatively cheap general-purpose battery applied to African winch-operators by Mkele (48), only produced an increase in multiple R from 0.66 to 0.74.

As practical abilities tend to be very specific, Drake (25) and others have maintained that the only way of covering them adequately is by means of tests that are close behavioural representations of the actual tasks the candidate will be required to perform on the job. Very probably, apparatus tests specifically designed for the above occupations, would have given better results, but in most of these projects all-round expediency would not have warranted such a refinement. With regard to other selection work done in connection with Air-pilotage, Air-traffic control, and machine-operation in the clothing industry, where there was sufficient economic or other justification for developing and applying special apparatus tests, such tests have definitely proved useful.

The more closely a test reproduces the criterion task in psychologically significant detail, the better its predictive value, but the more elaborate , costly and inconvenient it tends to become, until a stage is reached when there seems no point in having a separate test situation - the candidate might just as well be tested through the medium of a probationary period at the actual job. The latter procedure, which was extensively followed from the dawn of occupational specialization until well into the present century, might still be quite generally acceptable, were it not for the unfortunate fact that, because of the vast increase in the technical, behavioural, and economic complexity of human enterprise, it often means excessive material expenditure and disruption of productive work, sometimes accompanied by bodily hazard to candidates and others. Synthetic testing situations are therefore frequently necessary, and although never ideal, can fulfil a useful function. An example of one of these, carried to about the practicable limit of elaboration, is the "Link-trainer" type of test used in the Royal Canadian Air Force, which closely reproduces many facets of the air-pilot's task and has shown remarkable predictive value for pilot selection.

To the research psychologist, the main drawback of these complex tests is that they are almost as difficult to control and analyse into behavioural components, as the actual job situations. To the personnel selector this is of minor importance; provided the test, as a whole, has predictive usefulness, it does not really matter if the various factors contributing to the validity are numerous and so inextricably interwoven that they baffle identification. However, the personnel selector is ordinarily interested in improving his instruments, and to do this it is not sufficient for him to know that a test "works", but also, to some extent, how it works, which can be quite a problem if the test involves lots of variates. A precept attributed to the great experimentalist Faraday is worth remembering: "If you are going to ask Nature a question, let it be a simple one."

An alternative approach to the use of complex "job" tests may lie in the development of more general-purpose instruments, where the variance depends on factors rather than specific abilities, for, as Thorndike (69) has pointed out, complex "job-sample" tests overlap not only in their valid variance but also in their invalid variance for each job. This approach was extensively tried out in the American Air Forces, but was only successful when the occupation itself was well loaded with some factor or factors, as, for example, navigation, with numerical and general-reasoning factors. The more diversified occupation of the air-pilot was always covered better by complex than factor-pure tests. In this connection, Guilford (35; p.879) remarks : "In the search for valid pure tests, one finding is disturbing to the investigator who, following the traditional teachings on test construction, works toward maximizing the validity of each test. If the latter is the sole objective, we almost always end up with complex tests."

It may be practicable however, to develop a compromise in the form of somewhat restricted general-purpose tests to cover, say, the main sorts of sensory-motor skill required in a certain occupation or a few similar occupations, which would certainly be more expedient than constructing large numbers of specific apparatus tests. With something of this kind in view, Fleishman and Ellison (28) propose a more systematic enquiry into both tests and occupational requirements. Their factorial results have indicated that manipulative tasks generally involve four main factors, namely, manual dexterity, finger dexterity, speed of arm movement, and arm-hand steadiness. Therefore, a few apparatus tests adequately saturated with these factors, should afford a fairly wide occupational coverage.

Possibly there is often unwarranted concern over the fact that most existing general-purpose tests, while they tend to correlate quite highly with the theoretical aspects of practical occupations, are unrelated to the sensory-motor skills involved.

Tiffin (70) and others have stressed that the overall success or failure of the craftsman usually depends far more on factors of training, experience, mechanical comprehension, and judgment, than on basic dexterity. The performance, on a number of manipulative tests, of subjects who had had experience in various occupations, was investigated by Teegarden (68), whose results indicated that differences between occupations lay more in the ability to solve problems, react to a multiplicity of details, and make accurate movements, than in speed of hand movement or two-hand co-ordination.

Some practical tests, that have shown only very specific usefulness on normal subjects, can have considerable "general-purpose" value in assessing the vocational abilities of persons who are mentally or physically handicapped. Using a battery that was mainly practical, Murray (53) found that the potentialities of mentally retarded subjects for a wide variety of jobs, could be usefully assessed. A study on blind subjects by Hoffman (37) showed that their vocational earnings correlated significantly with their performance on the Minnesota Rate of Manipulation test and the Purdue Pegboard.

The uses of practical tests for gauging the vocational potentialities of unsophisticated normal subjects, such as tribal Africans, will be discussed further on.

Reliability and Validity

To be of use in occupational selection, any test, whether of the "printed" or "apparatus" type must fulfil two basic requirements - it must be both inherently reliable and specifically valid for the intended purpose.

In its psychometric sense, the term "reliability" implies the consistency with which a test functions. This is usually determined by repeating the test on the same group of subjects, or, after a single administration of the test, by comparing performance between the first and second half, or between several smaller sections. If a high degree of correlation is obtained between repetitions or sections, the test is regarded as reliable, which means, essentially, that chance plays only a small part in the production of the scores, and the individual differences in performance definitely reflect genuine differences in some ability or quality, at least within the confines of the particular test. The correlations produced by these procedures give what is called relative reliability, which is an adequate indication of reliability for most purposes. A more refined estimate, known as absolute reliability, is given by the standard error of measurement, which is an estimate of the extent to which the test scores obtained deviate from their "true" values, i.e. values that are free from chance factors and other errors of measurement. (Freeman, 29; p.78).

/ The term "validity" refers to the degree of precision with which a test measures the attribute(s) it purports to measure, and is reflected by the extent of correlation with some other known test of the attribute being measured, or with some criterion of achievement in an occupation. Although its meaning (in the psychometric sense) is very different from that of "reliability", it is actually a function of the latter, but the relationship is not reciprocal, for, while a grossly unreliable test cannot have any useful validity, a highly reliable test does not necessarily always have high validity - it can have any degree of validity. For example, the "Time" measure on the Variable Co-ordination test (23) which showed good inter-trial consistency on samples of both mine winch-drivers and air-pilots, was found to be significantly related to success in air-pilotage, but not to success in winch-driving. Reliability is intrinsic to a test, but not validity, for the test has a particular validity in relation to every criterion measure with which it may be compared. To quote Bingham (15; p. 221) : a test "... is valid <u>for the purpose in hand</u>. A test can be "good" only in relation to the specific purpose to which it is put." Gellerman (30) has also emphasized the relative nature of validity.

It is generally conceded that a test-measure with a reliability coefficient of 0.80 or larger can be used with confidence, but there are occasions when the validity may be sufficiently high to warrant the acceptance of a much lower reliability coefficient. A measure with a reliability of 0.60 and a validity of 0.50 would usually be far more useful for occupational selection purposes than one having, say, a reliability of 0.98 and a validity of 0.20. A very important consideration is the degree of reliability of the criterion-measure. If this is 0.60 it is still possible for a test with a reliability of as low as 0.40 to be quite a useful predictive instrument having a validity of 0.49. Ghiselli and Brown (31; p. 139) give a table showing the maximum validity coefficients possible with various combinations of test and criterion reliability.

Guilford (35; p.878) quotes some actual examples of unusually low test reliability accompanied by quite good validity. One of these, a 15 item judgment test in the Air Corps Qualifying Examination, had a reliability of 0.36 (odd-even) and a validity of 0.36. He observes ".... it looks as though the common advice that if a test shows validity one can forget about its reliability, might be sound, at least in some tests."

Most apparatus tests, designed and applied with reasonable forethought, show acceptable reliability. When two measures of performance are afforded, namely, time taken to perform the task and amount of error made, as in Steadiness (22) and Variable Co-ordination (23), the former is usually the more reliable, because amount of error, in terms of time off the correct track, is made up of intermittent occurrences that together constitute a time period far less than that spent on performing the task and chance factors have therefore a greater influence on this measure. A similar result is observable in a test like Dots-cancellation where the two measures are speed of work and the actual number of errors made. The continuous nature and longer duration of the former, favours consistency, whereas the intermittent nature and comparatively short duration of the latter makes it more subject to random influences that impair consistency.

This effect is not restricted to practical tests or tests affording measures of time and error. It is a basic psychometric rule, that within certain bounds, the "longer" the measure, the better its reliability. Prolongation need not necessarily be continuous, but can also be obtained by repetition. Collectively, several separate trials on, say, a short tracking task can be as dependable as a single trial on a long task. An illustration of the result of repetition in a perceptual task was given when two distinct techniques were used to obtain measures of Speed of perception and Span of attention (24). In measuring Speed of perception, a moving stimulus was exposed repeatedly at reduced speeds until the subject perceived it; in measuring Span of attention, a stationary stimulus was exposed only once for a brief period. The first test proved to have superior reliability.

Very high reliability coefficients resulting from inter-trial correlation, should be accepted with caution, because there is a likelihood that they have been spuriously augmented by some artifact in the particular test procedure employed. Even in tests of quite short duration, exceptionally high inter-trial correlations often result when the method of limits is used, because of the inherent "practice and expectancy" element which causes the response to a present stimulus to be influenced by that made to a preceding one. The subject's critical appraisal of individual stimuli is thus coloured by a tendency to respond consistently to the whole series, which produces a reliability coefficient that is an exaggerated representation of true test consistence. As Reuning (60) has pointed out, this "practice and expectancy" element characteristic of the method of limits, is probably responsible for boosting the reliability coefficients obtained on tests of Flicker-fusion threshold when this method is used. However, he has established by complete re-testing, which is the ultimate "proof of the pudding" in such cases, that the CFF measure has good actual consistence.

In test design, it is far easier to obtain reliability than validity. Much useful basic research can be done within the boundaries of a reliable test irrespective of the test's relationship to an outside criterion. For example, restricting an investigation to the small universe of one test such as Mirror-drawing or Dots-cancellation, it is possible to study, among other things, the structure of work curves, the relationship between speed and accuracy of performance, the effects of incentive, distraction, or fatigue, and differences between sexes, age groups, culture groups, or normal and psychopathic subjects. But when a test is to be used in selecting candidates for an occupation, unless it has some validity, as shown by a significant connection with success in that occupation, it has no usefulness at all, however reliable it may be.

One of the main reasons why validity is more difficult to achieve is that the criterion-measure must also be consistent and "genuine", but this consummation, though devoutly to be wished, is often not possible. The investigator seldom has direct control over the criterion-measure, and even when he has, the greater complexity of this measure as compared with the test, makes it much less amenable to control. Criteria of occupational success in the form of subjective assessments of the candidates by superiors or peers are commonly very unreliable; and even more objective measures like actual productivity on the job, promotion, bonuses, over-time earnings, absenteeism, accident-proneness, and achievement on trade-tests, although they may be influenced less by random factors, are often inadequate or not available. The least unsatisfactory criterion is provided by examination results after a course of training in the occupation. Here, however, there may be some tendency for the validity to be boosted artificially if the particular technique used in the examination bears a very close resemblance to that used in the test. It is a fact well known to psychometrists that the correlation obtained between written tests is not entirely attributable to the <u>matter</u> contained in the tests, but also, to some extent, to the similarity in external form - they have the common element of requiring the subject to identify printed signs on paper and inscribe responses on paper with a pencil. A simple perceptual-motor task like Dots-cancellation, which appears to involve very little intellectual ability, nevertheless shows some correlation with high-level intellectual tasks, probably because of the common pencil-and-paper medium. When, in addition to this common medium, the test and examination have also a similarity in structure, both requiring, say, a question and answer type of response, there may be a further artificial boost of validity. The test may, therefore, be rather less a predictor of competence in the examination itself than a predictor of ability to succeed in writing an examination. For selection purposes, this would not matter if the examination was always an accurate representation of all-round occupational competency, but unfortunately this requirement is rarely met, because the relationship between them is usually subject to considerable error variance.

Practical tests, in which the emphasis is on perceptual and motor response, ordinarily relate to such a restricted aspect of competence in a total occupation or a general examination purporting to represent true competence, that there is little likelihood of their validity being spuriously augmented to any appreciable extent, and the same would hold for pencil-and-paper tests that are highly specific. Even in the case of pencil-and-paper tests of the more general-purpose type, this augmentation of their validity is really insignificant when the criterion-examination comprises both written theory and practical performance in a variety of physical situations, which is the kind of examination given to air-pilots.

Before a test is used for selection purposes, it is advisable to do a validation study to ascertain empirically whether it does, in fact, relate to the criterion. Because this is inevitably a laborious and time-consuming procedure, there may be a strong temptation to apply the test directly for selection, on the rationale that it looks as though it incorporates something that is important in the job performance. Although there are occasions when practical exigencies warrant this procedure, it is always risky, for appearances can be misleading. An attempt by the writer to combine certain elements of "Steadiness" and Two-hand co-ordination into a single apparatus (Handlebars) did not produce a useful instrument for the prediction of air-pilot success, although Steadiness and Two-hand co-ordination measured on separate tests had proved of use for this purpose. In the Hand-foot reaction test, the use of a light stimulus to confirm correct response, on the assumption that visual stimuli are generally more important than auditory to the air-pilot, proved to be less successful than that of an auditory stimulus. The test of Dynamic anticipatory reaction, despite its promising appearance, was found to have no validity whatever for selecting mine winch-drivers. An external resemblance between test and occupation is no guarantee that the test will sample behavioural essentials of the occupation.

The uncertainty of relying on "assumed" or "face" validity also applies to pencil-and-paper tests. It is mentioned by Bingham (15; p.9) that there is at least one case on record in which a well-known test designed to measure mechanical ability gave better correlations with success in office work than certain tests specially designed to measure clerical aptitude; and another, in which an equally well known number- and word-checking test intended primarily for clerical workers, was found to have more validity for success in tool-making apprenticeship than the above test of mechanical ability. Vernon (74; p.105) notes that a mechanical assembly test used in the British army during the last war was more valid for predicting success in gross physical occupations than in skilled mechanical ones.

The thorny paths of personnel selection are strewn with many instances of "assumed" or "face" validity that appeared so promising at first, but had later to be rejected as counterfeit.

A validity co-efficient obtained with a particular test must always be interpreted in relation to the kind of sample from which it has emerged. Homogeneity in the sample structure means greater "curtailment" or "restriction of range", and in order to show validity the test has to discriminate very accurately between individuals. Heterogeneity means a comparatively wide dispersion of the performance-range between individuals, and a much coarser degree of discrimination by the test is adequate for it to show validity. Suppose, for example, that a test is applied to two distinct groups of pupils starting on an engineering course. The first consists entirely of men 18 to 20 years old who had all passed in science and mathematics in matriculation, and the second, of men and women varying greatly in age, educational status, subjects studied, and so on - in effect, a pretty random cross-section of adult human population. Later a correlation is done between test scores and the degree of success achieved on the engineering course. The coefficient obtained will be higher for the second "unselected" group, than for the "selected" group. Homogeneity of sample tends to depress validity and heterogeneity, to augment it; and the more intensive the selection, the greater becomes this effect of curtailment.

In establishing the validities of new apparatus tests for air-pilot selection in the S.A.A.F. after World War II, practical considerations always imposed the necessity of doing so on pretty highly selected groups. In addition to conforming to certain standards of age, education, and physical fitness, all candidates were screened on the basis of a battery of written tests such as Mental Alertness, Mathematical ability, Arithmetic, Technical and scientific reading comprehension and knowledge. Only the good performers on this screening battery, usually about a third of the initial sample, proceeded to the stage of doing the standardized battery of apparatus tests, comprising Air-pilots' arm-leg co-ordination, Hand-foot reaction, Steadiness, and Two-hand co-ordination (Moede-type). The experimental apparatus to be validated was also given at this stage, but the scores were not taken into account in subsequent assessment. Performance on the standard apparatus battery, and the personal interview of each candidate, first, by a trained personnel officer, and later by a board of expert Air-force officers, increased the selectivity still further. The group accepted to undergo flying training was therefore made up largely of the "cream" of the original sample, and to show validity against a pass/fail or rank-order criterion, the experimental test had to be capable of discriminating not merely between good, middling, and bad candidates, but within a restricted group of good ones. In view of this, it is hardly surprising that the apparent validity coefficients obtained with such experimental tests were often on the low side.

The higher the degree of correlation between the experimental test and one or more tests on which the selection was done, the greater is this effect of curtailment in depressing the validity coefficient of the experimental test. This applies particularly when an attempt is made to check the validity of a test which has actually constituted part of the battery on which the selection has been done. For example, when the test of Flicker-fusion threshold, which showed substantial experimental validity, was re-validated on subjects selected partly on the basis of their scores in this test, the resultant coefficients were very small.

When the experimental test is largely unrelated to any of the procedures that have been used in selecting the group on which it will be validated, the depressive effect of the curtailment on the coefficient is much less. This is probably a contributory reason why the test of Flicker-fusion threshold, which had no significant correlation with any written or apparatus test in the air-pilot battery, showed such a substantial experimental relationship to the flying training criterion.

The long term predictive validity of tests can be greatly depressed through changes in criteria with lapse of time.

Bass (5) mentions an instance where the successive merit ratings on a group of salesmen correlated 0.62 after an interval of 6 months, but only 0.29 after 42 months, and there was likewise a reduction in the correspondence between actual and predicted occupational success. The Steadiness Test, which had validity for the prediction of air-pilot success during the war, had practically none after the war, largely because in certain respects the training criterion had become less stringent.

Validity coefficients are also depressed through the use of global criteria like overall assessments of job efficiency or "pass/fail" on a flying course. As Biesheuvel (14; p.297) notes "Only an averaging effect can result from attempts to predict a criterion treated as if it were unitary when in fact it is composed of many things. No amount of refinement on the statistical side can overcome this basic weakness." While the global criterion may well be the most important, where possible, it is also advisable to investigate more specific sub-criteria, as was done in validating the test of Paired-associate learning, which showed no correlation with overall success or failure on a flying course, but it was found to relate significantly to certain branches of the course.

This principle of differential validation is also applicable in the test itself when more than one measure of performance is afforded, such as Time and Error, because often it is a particular measure that relates to the criterion; a composite score may dilute or even completely neutralize the relationship.

In view of the wide variety of influences operative in the production of a validity coefficient, some having an enhancing and some a depressing effect, it is not surprising that there is often considerable disparity between results obtained on different samples. The actual size of the coefficient is not the only factor to be taken into account in assessing the predictive usefulness of a test. Even a small degree of validity, provided it is significant, can justify the inclusion of the test in a selection battery, if the test measures some function not covered by the other tests. Of the various tests developed in the U.S.A. during the last war for selecting air-pilot candidates (Guilford and Lacey, 35; Melton, 46) approximately 60 were of the "pencil-and-paper" sort covering many qualities considered important in the aviator's vocation, and six were individual sensory-motor tests using apparatus. The best of the pencil-and-paper tests gave a multiple correlation of about 0.50 with success in flying training. When the six sensory-motor tests were included with these, the multiple correlation was increased to 0.70, although the median validity coefficients of the sensory-motor tests were generally low, being as follows : Complex Co-ordinator, 0.37; Two-hand Co-ordinator, 0.34; Rotary Pursuit, 0.27; Discrimination Reaction Time, 0.22; Finger Dexterity, 0.11; Steadiness, 0.09. Each of these tests made a useful contribution because it tapped a rather unique function; the mean intercorrelation between the tests being of the order of 0.20 (Range, 0.45 to 0.11).

When a choice has to be made between two "similar" tests, one of which has shown some validity for the particular predictive purpose, and the other none, the former will have priority; and similarly, when the one has shown an exceptionally high validity, and the other a just acceptable validity. In cases where the difference between the respective validities is not so obviously great, it can be difficult and frequently impossible to judge the relative merits of the two tests for the purpose in hand, because unless the samples are very large, or the validity coefficients represent the results of repeated applications of the tests, the difference between the coefficients cannot be shown to be statistically significant. Since the last war, validation studies on tests for pilots in the S.A.A.F. have had to be restricted to rather small groups, and except in those cases where the difference between validites is obviously large, there is no empirical justification for concluding that one measure is superior to another as a predictor of air-pilot success. In such instances, when objective data by itself is not very helpful, it is quite permissible to give preference to certain tests on rational grounds - in other words, to take "assumed" or "face" validity into account. As Bingham (15; p.223) puts it : "The psychologist either has to let his client wait for months or years while he resorts to new experiments and elaborate statistical analyses in order to compare the merits of similar tests, or else he has to lean heavily, as a physician does, on his informed common sense." Purely utilitarian considerations also carry weight in making a decision. If either of two tests of say, co-ordination, would make about the same contribution to a battery for a certain purpose, preference will be given to the one that is more convenient and economical to construct and apply.

Ultimately, the inclusion of certain tests in a selection battery depends on overall expediency. For example, in developing a battery for the selection of mine winch-operators, four elementary screening tests that could conveniently be administered as group tests, and had proved of value in the general classification of African mine-workers, were supplemented by four sensory-motor tests requiring individual administration, namely, Two-hand Pursuit, Dynamic Coincident Reaction, Code-signal Reaction, and Two-hand Co-ordination (Moede type). The simple screening battery by itself, consisting of Cube Construction, Letter-and-numeral Sorting, Object-Sorting, and Tripod-Assembly, gave a multiple correlation of 0.66 with success in winch-driving; the sensory-motor tests, as a group, 0.71; and the entire battery, 0.74. Although the addition of the sensory-motor tests produced an appreciable gain in the validity of the battery, in terms of "operating characteristics" (Mkele, 48) it was still not regarded as sufficient to justify the inclusion of these tests, because of the time and expense involved in their construction and administration.

Such a conclusion may be warranted in relation to an occupation like mine winch-driving, where losses sustained in training operators who turn out failures are not very great, and where serious damage to equipment or physical injury resulting through operators' incompetence, is fortunately of comparatively rare occurrence. But in selecting for a hazardous and highly-skilled occupation like air-pilotage, where training is extremely expensive, and accidents often mean loss of life and great destruction of equipment, a gain in battery validity, equivalent to the above would be very valuable, and fully justify the retention of the more costly and time-consuming tests.

In the foregoing discussion, the general term "Validity" was used to designate the usefulness of a test for forecasting behaviour or achievement in some other sphere of activity. This is also known more particularly as Predictive Validity. There are other concepts of validity which can be described briefly as follows.

Operational Validity implies that the test provides an acceptable medium for measuring and assessing activities of a definite kind. For example, the Moede-type Two-hand Co-ordination machine tests a certain pattern of muscular co-ordination, but not all muscular co-ordination. As this test has been found to discriminate well between individuals, with respect to the particular co-ordinative activities involved, it can be said to have "Operational Validity", which is closely related to the concept of 'Reliability".

Tests that are operationally valid may or may not be predictively valid. For example, tests of Steadiness which have been found to possess "Operational Validity" as measures of fine muscular-control, will probably be of very little use in predicting success in surgery, because, although steadiness plays a part in such a vocation its overall importance is outweighed by that of other factors such as knowledge, experience and judgment. On the other hand, Steadiness tests have "Predictive Validity" for success in rifle-marksmanship, where fine muscular control is a principal requirement. (62, 66).

Content Validity has some similarity to "face validity", in so far as its estimation is hypothetical rather than empirical. It refers to the extent to which the test is representative of the activities it is desired to measure, in accordance with the best subjective judgment available, sometimes, but not always, aided by previous empirical findings. While this concept is usually confined to printed tests composed of a large number of individual items, which should all bear some relevance to the variable to be measured, it can also apply to apparatus tests. As an instance, might be mentioned, that in planning the specific

Arm-leg Co-ordination test for air-pilots, it was necessary to weigh the merits of various alternative mechanical systems and even build a few crude experimental lay-outs, before deciding on a system that appeared to provide the most promising task for measuring the kind of co-ordination required by the air-pilot. To have the best Content Validity this task had to reproduce, as far as possible, the essential pattern of co-ordination in piloting an aircraft, rather than that involved in some other activity like riding a bicycle, driving a motor car, or rowing a boat. Good content validity is the essence of a good specific-duty test, whether of the printed or apparatus kind.

Concurrent Validity is a rather new concept which appears to have both a general and a more specific meaning. Generally it refers to the degree of relationship between test performance and the status or classification of the subjects at the time they were tested. For example, in validating general-purpose tests for African mine-workers, men from three different levels in the occupational hierarchy on the mine were tested, and most of the tests proved to be concurrently valid in that there was a significant correspondence between achievement level in the tests and the occupational level at which the subjects were being employed. Similarly, the Mark I model of the Air-pilot's Arm-leg Co-ordination Test was applied to both a group of experienced pupil-pilots and a group of inexperienced recruits, the concurrent validity of the test being indicated by the fact that the former group performed significantly better.

More specifically, Concurrent Validity refers to the usefulness of a test as a substitute for some already existing procedure measuring similar abilities or traits. In this case, the existing procedure is the criterion. From the standpoint of personnel selection, it is usually only expedient to seek the establishment of validity in this specific sense if the existing procedure is costly or cumbersome and the new one promises distinct economic or other practical advantages; for example, with the object of replacing an elaborate apparatus test which has well established predictive validity, with a behaviourally similar, but constructionally, much simpler piece of apparatus, or even a pencil-and-paper substitute. If the new procedure correlates highly with the old, it is regarded as being concurrently valid.

Construct Validity is a new and more refined conception of "Content Validity"; so refined, in fact, that its precise meaning (if it has any) is probably only comprehensible to the esoteric few who (a) have immersed themselves long and deeply in the copious controversial literature, and (b) have a mental-alertness level well above the cut-off point for admission to membership of "Mensa".

Among other things, the concept would seem to imply that: (1) The hypothesis underlying the design of the test is based extensively on empirical findings from previous tests of the quality or attribute to be measured. (2) The quality or attribute has been analysed and defined (by factor analysis or other methods). (c) Through adequate experimental application, proper scales of measurement have been established whereby the performance level or "status" of a single individual can be fairly accurately assessed in terms of the specified quality or attribute, and in relation to the range of measurements derived from the population samples that have been tested experimentally. In general, the emphasis appears to be on empirical rather than hypothetical determinations of test content. The Englishes' dictionary (26) states that the "construct" must be explicit : "Thus, for a test of manual skill, one needs to know exactly what kinds of behaviour exemplify manual skill".

According to Cronbach (19; p.120), "Construct validation is an analysis of the meaning of test scores in terms of psychological concepts." Freeman (29; p.99) defines it thus: "A test's construct validity (including factorial validity) will indicate the psychological operations on the basis of which one's test performance may be explained and evaluated". Loevinger (43; p.636) maintains that "....since predictive, and content validities are all essentially ad hoc, construct validity is the whole of validity from a scientific point of view". Cronbach and Meehl (18; p.282) say that "Construct validation is involved whenever a test is to be interpreted as a measure of some attribute or quality which is not 'operationally defined'. The problem faced by the investigator is, 'What constructs account for variance in test performance?' " According to the APA (1; p.14) "Construct validity is ordinarily studied when the tester has no definite criterion measure of the quality with which he is concerned, and must use indirect measures. Here the trait or quality underlying the test is of central importance, rather than either the test behaviour or the scores on the criteria."

Obviously, between some of the above interpretations, there are considerable differences. One gets the impression that this concept is less a concept of "validity" than of methodology, and as such, it lends itself to various interpretations, all of which may be scientifically acceptable, although they place the emphasis rather differently. In the opinion of Bechtoldt

(7; p.628), "The renaming of the process of building a theory of behavior by the new term "construct validity" contributes nothing to the understanding of the process, nor to the usefulness of the concepts."

<u>Cross Validation</u> simply means a repetition of the process of predictive validation with one or more groups of other subjects, which provides a truer estimate by tending to even out chance factors that may have induced a spurious augmentation or depression of the preceding correlation(s). To establish with reasonable certainty that the predictive validity of a test is dependable enough for selection purposes, more than one validation study is advisable, particularly when the test samples are on the small side. In this connection, Biesheuvel (14; p.317) has noted : "The still far too prevalent practice of equating unreplicated statistical significance with practical usefulness invites one-sided attacks on the scientific status of personnel selection."

It has been pointed out by Freeman (29; p.99) that content, construct and concurrent validities are actually "evaluations of the extent to which the device estimates an individual's status at the time the test was administered. From the viewpoint of applied psychology, every test, whatever the type, must ultimately be shown to have predictive validity." He considers that the latter is the most important characteristic of a test.

Affective, Conative and Social Aspects

The application of apparatus tests for objectively measuring traits of temperament and personality, is comparatively rare, mainly because so little has been achieved in developing dependable tests of this kind. Among the more successful are electrical skin-conductance, alpha-frequency on the electroencephalogram, involuntary hand-tremor, and threshold for flicker-fusion. Apparently the latter provides some measure of "primary-secondary" function that seems to reflect a temperamental quality associated with success or failure in learning to pilot aircraft.

Although objective measurement should be used in scientific work wherever possible, subjective observation can have great value, particularly in biological fields, and in many situations it is the only practicable method. As a rule, scientific achievement of any worth has always resulted from the combination of both systems. Temperament and personality qualities have been successfully assessed by "projective techniques" like word-association, sentence-completion, and thematic apperception, in which the interpretation of results is a largely subjective matter. However, for purposes of occupational selection, the majority of both these tests and the more objective "physiological" procedures mentioned above, have the drawback of taking rather long to administer and interpret, and requiring the services of highly specialized testers.

When a subject performs a practical test, there are often facets of behaviour such as talkativeness and the kind of question asked, distractibility, restlessness, manner of setting about the task, overt signs of tension or fear, reactions when mistakes are made, perseverance, confidence, interest, willingness to co-operate, and so on, giving a trained observer supplementary information, that, for selection purposes can sometimes prove of considerably greater worth than the objective measure of the candidate's proficiency at the task. For although the latter may have substantial predictive usefulness, this usually relates to a restricted aspect of the occupation, whereas suitability of temperament, personality, and interests is absolutely fundamental for successful adjustment to the occupation as a whole. It is probable that in all fields of human activity, temperament and personality variables play a major part in the causation of accidents, from minor ones to the most serious.

The following apparatus tests were included by Reuning and Rosen (58) in their battery when they investigated the observation of temperament variables in skill-test situations: Mechanical aptitude, Hand-foot reaction; Steadiness; and Floating-effect (co-ordination). While a certain lack of relationship between the ratings by three different pairs of observers was found on individual items, the overall assessments appeared, for the most part, to be reasonably consistent.

In assessing the temperamental aspects of a subject's behaviour on sensory-motor and other performance tests, there are two main forms of bias to be considered. The first relates to the actual proficiency of the subject at the task. Following on their previous study, Reuning and Rosen (59) reported a consistent tendency among observers to rate the more successful subjects highly on temperament traits regarded as desirable. The second bias relates to the observer, who, as Baehr (4) has demonstrated, is influenced in his judgment by his own temperament make-up, and tends to be more favourably disposed towards certain subjects than others.

Notwithstanding its limitations, the observational method was applied with considerable success in the Aptitude Tests Section of the S.A.A.F. during World War II. Biesheuvel (9) found it very useful, not only in refining the selection of pilots, navigators, and air-gunners, but also in discriminating between fighter- and bomber- or reconnaissance-pilots. Objective test measures alone proved much less successful than the combination of these with critical observations of overt behaviour in the test situation.

While many practical tests provide situations for observing affective, conative and social aspects of behaviour, the most useful in this respect are those well loaded with a frustration element, like Chopsticks (21), or some other stress such as the threat of electric shock in Variable Co-ordination (23) and Steadiness (22).

Although it is somewhat less specialized than the "projective technique" or certain of the more complex "physiological" procedures, the observational study of overt reaction in performance tests does require considerable experience and judgment on the part of the tester if it is to be of any value in helping to sound the occupational suitability of candidates. Not only may certain significant "clues" be missed, but others may be quite wrongly interpreted. Consider, for example, manifestations of "tension" by the subject. A critical appraisal of several early investigations on emotional reaction to stress situations by means of practical tests was done by Arnold (3), who, on the strength of his own experimental results and those of others, concluded that there are two main kinds of tension, voluntary, which is the result of conscious effort by the subject and actually assists his performance, and involuntary, which is neurotic in origin and tends to inhibit performance. This conforms to our observations in testing air-pilot candidates on various practical tests. Tension on the controls, unless it is accompanied by very marked tremor of the limbs, is not necessarily an unfavourable sign. On the Variable Co-ordination test (23), objective confirmation of this was afforded by a significant positive correlation obtained between heavy, prolonged pressure on the foot-controls and subsequent success in learning to fly an air-craft.

If it is practicable to include a "projective" or "physiological" test of temperament qualities in a selection battery, so much the better, but they should not be regarded as substitutes for the careful and systematic observation of subjects' general behaviour in sensory-motor and other performance tests, which are dynamic situations favouring the manifestation of important idiosyncrasies that are likely to remain obscured in the more physically static type of test situation.

Skawran (65) employed the Willemse-board as a medium for assessing aspiration level. A long, sloping board had to be manipulated by the subject to steer a rolling marble into a pocket that could be fixed at various distances along the edge of the board, in accordance with his stated opinion of the distance he felt he could achieve. Significant correlations were obtained between certain measures of aspiration and success in air-pilot training.

The situational test can also be of value in providing global estimates of an ability. Kamfer (41) used a piece of meteorological apparatus to assess the instructional ability of potential army instructors. The purpose and functioning of the apparatus was first explained to each candidate, and he was then required to describe these to a group of assessors. Ratings were found to correlate significantly with examination grades achieved by the candidates at the Military College.

Application to African Subjects

A field of personnel selection in which apparatus tests are proving very useful, is that concerned with the testing of Africans for educational and vocational placement. Many candidates lack the minimal degree of literacy and sophistication required for doing pencil-and-paper or pictorial tests, and even to the more "advanced" African, such tests are often "culturally inappropriate". As Biesheuvel (11) observes : "His score may be low, not because the particular act of reasoning or abstraction required by the problem is beyond him, but because the assignment has been put to him through the medium of a foreign symbolism which he does not understand." Practical tests can, of course, also be culturally inappropriate for Africans, but usually they are less defective in this respect than printed tests, because adjustment to physical reality, and the handling and arranging of concrete objects, are basic processes common to all cultures, although they may differ in specific detail.

The experiences of European observers in Africa, including pioneers like Burton (16) and Speke (67), afford many instances showing that primitive Africans respond to the physical novelties of European civilization far more readily than they do to intellectual, social, moral and aesthetic concepts. A hundred years ago, Mtesa, the intelligent young king of Uganda soon mastered the basic technicalities of the most modern firearms of the period, unwisely introduced by Speke, and became a very enthusiastic and quite passable marksman, but he could never realize that by English standards it was unsporting to shoot tame domestic cows at point-blank range merely to show off before his harem, or that there was anything wrong in telling fibs, breaking promises, purloining the property of others, and cruelly executing his subjects for trifling breaches of court etiquette. It was also quite beyond his comprehension that a European should value personal privacy, appreciate landscape for its own sake, and be unimpressed by the beauty of aristocratic Waganda females so fat that they could not stand upright.

Although the spread of European culture in some parts of Africa has induced many radical changes in the attitudes of the indigenous peoples to European values, adaptation continues to be more positive in the sphere of material things and physical activities than in that of abstract ideas, and it is usually through the medium of practical situations, involving apparatus of some kind, that the tribal African's ability can be most fairly assessed. This would also apply to the less sophisticated inhabitants of other parts of the world who are likewise still in the process of adapting to a more complex way of life forced upon them by the spread of Western or Neo-eastern "civilization".

The nuisance of having to rely so extensively on apparatus tests in African selection, is considerably mitigated by the fact that tests of this kind usually have greater "general purpose" versatility for Africans than for Europeans. One probable reason for this is reflected in a view expressed by Murray (52; p.73) that there is "a strong manipulative element in the structure of African intelligence". Another, is that African population samples are, for the most part, very heterogeneous, because there is very little automatic pre-selection.

Some of the apparatus tests originally developed in this Institute for the classification of African labour in the South

African gold-mining industry, have also proved of value in various secondary industries like motor-car assembly, building construction, brick-and-tile making, and the manufacture of metal goods. In West Africa, they have been successfully applied on both gold-mining personnel and military artisans; and in Kenya, for the selection of workers in the tobacco industry. Another example is provided by two tests of implement manipulation (21), namely, Chopsticks and Tweezer-nozzle, which were originally intended by the designer for the measurement of quite fine motor skills, and yet, when applied by Hudson, Mokoatle and Mbau (39) on a sample of African operatives in a factory producing nuts and bolts, were found to correlate with success at a comparatively crude assembly task like "nutting". In another organization, specializing in the production and assembly of steel pipes and fittings, the Tweezer-nozzle test is contributing usefully to a battery for the occupational grading of African personnel, and here, too, there is very little external resemblance between the kind of manual skill measured by the test, and that required in the job situations.

It seems, therefore, as though some tests which are mainly limited to the tapping of rather specific manual skills in European subjects, when administered to Africans, also provide significant measures of cognitive ability and, what Vernon (74) calls "general adaptability to the unfamiliar testing situation". While there is considerable evidence, provided by Murray (52) and others, that the structure of the African mind is characterized by factorial simplicity, this has usually been derived from a study of African performance in "European" tests. Whether it also applies in the African's own cultural fields of arts and crafts, language, music, intellectual games, folk-lore, and adjustment to natural environment, has still to be demonstrated. In any test situation devised and administered on conventional European lines, the African is handicapped by the additional burden of having to contend with a strange milieu, and the less contact he has had with European culture, the greater this handicap. In validating apparatus tests for the screening of African labour on the Witwatersrand gold mines, Hudson (38), Gouws (33) and Biesheuvel (10) found that the subjects' educational level and length of mining experience (which can be regarded as quite good reflections of the amount of previous meaningful association with European culture), had a marked influence on certain test performance. Gouws (34) has also reported differences in the test performance of respective tribal groups. Subjects from the East Coast and the Cape Province tended to be superior to those from other South African provinces and the Tropics. The fact that Mocambique and the Cape of Good Hope have had longer exposure to European influence than the other parts of Southern Africa, may be significant. Biesheuvel considers, however, that regional differences are mainly due to differential selection.

It is probably because of qualitative rather than quantitative differences between the mental structure of the respective groups, that on some European-conceived tests the mean scores of Africans tend to be lower than those of Europeans. As a rather extreme hypothetical example of the converse situation, imagine a group of typical European city-dwellers being "tested" in the Kalahari by the Bushmen to determine their ability at tracking game, finding food and water under the most adverse conditions, identifying numerous plants and insects, making and using bows and arrows, mastering the intricate details of traditional chants and dances, steering true courses across pathless wastes without the aid of a compass, and sustaining morale under the stress of thirst, hunger, fatigue, and extremes of heat and cold. Very few of the European candidates would show themselves to be "good occupational risks" in this strenuous and precarious way of life, and for most of the failures there would certainly be no further aptitude testing, either in the Kalahari or anywhere else.

A point which should always be kept in mind when selecting Africans, particularly those who have had little contact with European culture, is that test scores are likely to be an understatement of true ability. A candidate who fares indifferently on the tests may still succeed tolerably well on the job after a period of adaptation and training.

Personnel selectors would generally agree that whatever the race or degree of sophistication of the candidates, it is desirable for them to perceive some logical connection between the tests and the occupation. Motivation is thereby increased, and the test scores are likely to be a more accurate index of ability. This is another consideration favouring the use of practical situations in selecting for practical jobs. The candidate's idea of the connection between test and job need not necessarily be a true one. It can be quite imaginary, and provided it appears sound to him, still useful. In an interesting account of the reaction of African mine-workers to aptitude testing, Masilela (44) mentions some of their ideas about the relationship between the apparatus tests and the actual jobs underground. For the most part these were pretty close assessments of the functions of the tests, although the candidates founded their notions on the incorrect, but quite logical assumption, that the testing depot was a "school" and the tests were specific devices for training them to acquire proficiency in the various jobs.

In administering tests to Africans who have spent most of their formative years within their own tribal environment, observers are faced with the problem of communication. There are so many different languages and dialects that it would hardly be possible to evolve and apply standard test instructions in the home language of all candidates. A lingua franca like the Kiswahili of Equatorial East Africa or the Fanakalo of the South African gold mines has rather limited usefulness when the test sample is very polyglot. The use of apparatus tests enables practical demonstrations of test procedure to be given, which require little or no supplementary verbal comment. When testing large groups, the observer administers the mime in conjunction with a large-scale demonstration model of the apparatus, or projects a short motion-picture on a screen. The latter procedure, originally developed by Biesheuvel (10) and Hudson (38) for use on the Witwatersrand gold mines, has since been widely applied in African selection, and possesses the merit of not depending on subjective ability to demonstrate, which may vary with different observers, and also when exercised by the same observer on different occasions.

Pencil-and-paper tests, particularly if the material is of an abstract nature, are far more difficult to demonstrate through a non-verbal medium than simple apparatus tests, and the observer seldom feels certain that all the subjects know exactly what they are required to do. Dr. Reuning of this Institute has found that even in a simple pencil-and-paper test of "gestalt-continuation", which involves reproducing a rhythmic visual pattern by joining printed dots with short pencil lines, many illiterate subjects fail to grasp the instructions, which are given carefully in mime, and there is the further complication that to persons who are not used to handling a pencil, the joining of dots with pencil-strokes is quite an exacting task by itself. He proposes, in his next investigation, to modify the procedure by including a short period of pre-test training to ensure that the test requirements are understood and to give some elementary practice in the sensory-motor task of manipulating a pencil.

Conclusion

Whatever the ethnic composition or degree of sophistication of the sample on which selection is being done, and whether the main reliance is on "practical" or "theoretical" tests, it should never be forgotten that, firstly, personnel selection constitutes only one branch of the extensive and varied study of the relationship between man and his working environment, which Biesheuvel (12) aptly calls "occupational science"; and secondly, that the ability of a candidate to perform certain work is only one requirement for his success in an occupation. His temperament, personality, interests, values and attitudes are also important.

Modern industrial organization is becoming more and more aware that in repetitive production, particularly, the overall efficiency of the worker depends less on actual motor skill than on the satisfaction which he experiences from exercising such skill as he possesses, his relationships with co-workers and overseers, adaptation to the working environment as a whole, and last, but not necessarily least, the stability of his domestic life. Emphasis on the mechanization and rationalization of the work itself and on the technical efficiency of the worker as an animated machine, which was the credo of most Victorian industrialists, and even of some psychologists like Münsterberg (51) in a later era, is being tempered by more humanistic approaches, and by making the well-being of man, rather than his productivity, the end in view, his productivity is actually increased. As a rule, industry sustains far greater material losses through absenteeism, accidents, occupational psychoneurosis, labour-management disagreement, and other forms of maladjustment relating to social, temperamental, and personality factors, than through any sensory-motor or cognitive inefficiency of workers on the actual job.

Summary

1. Concrete things are more fundamental than abstractions, and most occupations are concerned with the direct or indirect manipulation of concrete things.

2. Tests used in personnel selection fall into two broad classes : (1) General-purpose, measuring the more basic attributes and applicable to many different occupations; and (2) Specific, measuring attributes largely unique to a certain kind of occupation. Either class can be subdivided into (a) Printed tests, and (b) Practical tests involving more concrete apparatus.

3. General-purpose tests are usually more economical and convenient than Specific tests, and printed tests, than apparatus tests. There are practical factors not sampled by printed tests.

4. As a rule the more successful and widely used apparatus tests are constructionally simple. The use of elaborate equipment is only worthwhile for high-level or vital selection, and even here, complication beyond a certain point ceases to be useful. Most apparatus tests have rather restricted application because practical abilities tend to be very specific.

5. Important elementary implications of reliability, validity, and the criterion measure, are briefly discussed. Practical

tests usually show acceptable reliability, "larger" measures being the more reliable. Sometimes reliability may be spuriously augmented by the particular test procedure. Validity, which is closely related to the internal consistency of the criterion-measure, is generally more difficult to achieve than reliability. "Assumed" or "face" validity can be deceptive. The size of a validity coefficient is not the only factor determining the value of a test. Ultimately, the value of a test is dictated by overall expediency.

6. Practical tests, particularly those having an element of stress, provide useful situations for assessing temperament and personality qualities by observation. However this subjective technique requires a certain modicum of experience and judgment.

7. Because orientation towards concrete reality is the greatest common factor pervading all cultures, practical tests are usually less "culturally inappropriate" for Africans than pencil-and-paper or pictorial tests. Nevertheless, in any test situation devised and administered on European lines, the African has to contend with a strange milieu. Practical tests have greater "general-purpose" value for Africans than for Europeans, mainly because of qualitative rather than quantitative differences in the mental structure of the respective groups. Practical tests usually give subjects a better idea of the connection between test and job, which favours motivation. They also have the advantage of being demonstrable to polyglot groups without the need for verbal explanation.

8. Personnel selection is only one branch of "occupational science", and the "ability" of a candidate is only one requirement for his success in an occupation. His temperament, personality, interests, values and attitudes are also important.

REFERENCES

1. AMERICAN PSYCHOLOGICAL ASSOCIATION, Committee on Psychological Tests. Technical recommendations for psychological tests and diagnostic techniques. Washington, D.C., APA, 1954.

2. ARBOUS, A.G. and H.S. SICHEL. On the economies of a pre-screening technique for aptitude test batteries. Psychometrika, 1952, 17, 331-346.

3. ARNOLD, M. A study of tension in relation to breakdown. J. gen. Psychol., 1942, 26, 316-346.

4. BAEHR, M.E. Problems in the N.I.P.R. technique of temperament assessment by means of observational methods. (Paper delivered to S.A. Association for Advancement of Science, Kimberley, July, 1949). Also in: Bull. Nat. Inst. Personnel Res., 1951, 3, No.3, 19-36.

5. BASS, B.M. Further evidence on the dynamic character of criteria. Personnel Psychol., 1962, 15, 93-97.

6. BECHTOLDT, Harold P. Selection. In: Stevens, S.S., ed. Handbook of Experimental Psychology, pp. 1237 - 1266. New York, Wiley, 1958.

7. BECHTOLDT, Harold P. Construct Validity : a critique. American Psychologist, 1959, 14, 619 - 629.

8. BECKMAN, R.O. A new scale for gauging occupational rank. <u>Personnel J.</u>, 1934, 13, 225 - 233.

9. BIESHEUVEL, S. An observational technique of temperament and personality assessment. (Paper delivered to S.A. Association for Advancement of Science, Kimberley, July, 1949). Also in: <u>Bull. Nat. Inst. Personnel Res.</u>, 1949, 1, No. 4, 9 - 27.

10. (Biesheuvel, S.) South African Council for Scientific and Industrial Research, National Institute for Personnel Research. <u>Aptitude tests for native labour on the Witwatersrand gold mines.</u> Johannesburg, N.I.P.R., 1949.

11. BIESHEUVEL, S. The study of African ability, Part II. <u>Afr. Stud.</u>, 1952, 11, 105 - 117.

12. BIESHEUVEL, S. The relationship between psychology and occupational science. <u>J. soc. Res.</u>, (Pretoria), 1954, 5, 129 - 140.

13. BIESHEUVEL, S. <u>The limitations of aptitude tests in selection for specific occupations.</u> Address delivered to the Annual Congress of the South African Psychological Association, Stellenbosch, July, 1955.

14. BIESHEUVEL, S. Personnel Selection. <u>In</u>: Farnsworth, P.R. <u>et al.</u>, eds. <u>Annual Review of Psychology</u>, Vol. 16, pp. 295 - 324, Palo Alto, California, Annual Reviews, Inc., 1965.

15. BINGHAM, W. van D. Aptitudes and Aptitude Testing, 11th ed. London, Harper, 1937.

16. BURTON, R.F. First Footsteps in East Africa. London, Dent and Dutton, 1910. (Original publication, 1856).

17. COX, J.W. Mechanical Aptitude, its Existence, Nature and Measurement. London, Methuen, 1928.

18. CRONBACH, L.J. and P.E. MEEHL. Construct validity in psychological tests. Psychol. Bull., 1955, 52, 281-302.

19. CRONBACH, L.J. Essentials of Psychological Testing. New York, Harper. 2nd Ed., 1960.

20. CRONBACH, L.J. and Goldine C. GLESER. Psychological Tests and Personnel Decisions. Urbana, Univ. of Illinois Press, 2nd Ed., 1965.

21. DE WET, D.R. Two tests of implement manipulation. J. Nat. Inst. Personnel Res., 1957, 7, 75-79.

22. DE WET, D.R. An improved steadiness apparatus and its validity for air-pilot selection. J. Nat. Inst. Personnel Res., 1960, 8, 122 - 136.

23. DE WET, D.R. A variable co-ordination test and its potentiality as a gauge of aptitude for airmanship. Psychol. Afr., 1962, 9, 86 - 99.

24. DE WET, D.R. Measures of speed of perception and span of attention on air-pilot candidates. Psychol. Afr., 1964, 10, 206 - 218.

25. DRAKE, C.A. New developments in worker analysis. Occupations, 1941, 19, 499-504.

26. ENGLISH, H.B. and Ava C. ENGLISH. A Comprehensive Dictionary of Psychological and Psychoanalytical Terms. New York. Longmans, Green, 1958. p.575.

27. FLEISHMAN, E.A. and W.E. HEMPEL. Factorial analysis of complex psychomotor performance and related skills. J. appl. Psychol., 1956, 40, 96 - 104.

28. FLEISHMAN, E.A. and G.S. ELLISON. A factor analysis of manipulative tests. J. appl. Psychol., 1962, 46, 96 - 105.

29. FREEMAN, Frank S. Theory and Practice of Psychological Testing. New York, Holt, Rinehart and Winston, 3rd Ed., 1962.

30. GELLERMAN, S.W. Personnel testing - what the critics overlook. Personnel, 1963, 40, 18 - 26.

31. GHISELLI, E.E. and C.W. BROWN. Personnel and Industrial Psychology, 2nd Ed. New York, McGraw-Hill, 1955.

32. GOTSMAN, E. A sequential procedure for selecting and classifying African industrial workers and labourers. J. Nat. Inst. Personnel Res., 1960, 8, 117 - 121.

33. GOUWS, D.J. The classification of native mine labour into job families. Bull. Nat. Inst. Personnel Res., 1950, 2, No. 2, 4 - 20.

34. GOUWS, D.J. The influence of tribal origin on the test performance of native mine labour. Bull. Nat. Inst. Personnel Res., 1950, 2, No. 2, 77 - 90.

35. GUILFORD, J.P. and J.I. LACEY (eds.). Printed Classification Tests, AAF Aviation Psychology Research Program Reports, No. 5, Washington, U.S. Govt. Printing Office, 1947.

36. HINES, Mildred and J. O'CONNOR. A measure of finger dexterity. Personnel J., 1926, 4, 379 - 382.

37. HOFFMAN, S. Some predictors of the manual work success of blind persons. Personnel guid. J., 1958, 36, 542 - 544.

38. HUDSON, W. The Occupational Classification of Africans. Ph. D. thesis. Johannesburg, Univ. of the Witwatersrand, 1953.

39. HUDSON, W., N. MOKOATLE and G.G. MBAU. The influence of training and practice on the test and work performance of a

small sample of African workers. J. Nat. Inst. Personnel Res., 1958, 7, 88 - 94.

40. HUMPHREYS, L.G., C.E. BUXTON and H.R. TAYLOR. Steadiness and rifle marksmanship. J. appl. Psychol., 1936, 20, 680 - 688.

41. KAMFER, L. Army instructor selection : development of a miniature situation test for the assessment of instructional ability and its contribution to a test battery. Psychol. Afr., 1962, 9, 44 - 59.

42. KEANE, F.L. and J. O'CONNOR. A measure of mechanical aptitude. Personnel J., 1927, 6, 15 - 24.

43. LOEVINGER, Jane. Objective tests as instruments of psychological theory. Psychol. Rep., 1957, 3, 635 - 694.

44. MASILELA, A.M. The reaction of African mineworkers to aptitude testing. Bull. Nat. Inst. Personnel Res., 1950, 2, No. 2, 21 - 23.

45. McMAHON, D. Selection and follow-up of engineering apprentices. Occup. Psychol., 1962, 36, 53 - 58.

46. MELTON, A.W. ed. Apparatus Tests. Army Air Forces Aviation Psychology Program Research Reports, No. 4. Washington, U.S. Govt. Printing Office, 1947.

47. MEREDITH, G.P. Models, meanings and men. In: Cohen, J. ed. Readings in Psychology, pp. 66 - 84. London, Allen and Unwin, 1964.

48. MKELE, N. Validation of aptitude tests for the selection of winch-drivers on the Witwatersrand gold mines. J. Nat. Inst. Personnel Res., 1953, 5, 100 - 109.

49. MOEDE, W. Arbeitstechnik. Stuttgart. Ferdinand Enka Verlag, 1935.

50. MONTGOMERY, G.W.G. Predicting success in engineering. Occup. Psychol., 1962, 36, 59 - 68.

51. MÜNSTERBERG, H. Psychology and Industrial Efficiency. New York, Houghton Mifflin, 1913.

52. MURRAY, C.O. The Structure of African Intelligence. A factorial study of the Abilities of Africans. M.A. thesis, Durban, Univ. of Natal, 1956.

53. MURRAY, Evelyn. Developing potential skills of the retarded. Emplmt. Secur. Rev., 1956, 23, 35 - 36.

54. PATERSON, D.G. et al. Minnesota Mechanical Ability Tests Minneapolis, Univ. of Minnesota Press, 1930.

55. PATIN, J. and S. NODIOT. Le test mécanique de P. Rennes. Bull. Cent. Étud. Rech. Psychotech., 1962, 11, 61 - 97.

56. RAMANATHAN, G. The digital correlate of mental functions. J. educ. Psychol., (Baroda), 1960, 18, 268 - 272.

57. REMMERS, H.H. and J.W. SCHELL. Testing the O'Connor Wiggly block test. Personnel J., 1933, 12, 155 - 159.

58. REUNING, H. and J. ROSEN. Analysis of items recorded by observers of behaviour in skill test situations. J. Nat. Inst. Personnel Res., 1953, 5, 36 - 56.

59. REUNING, H. and E.J. ROSEN. (From a report by D.R. Pitt and D. Rapozo). The effect of quality of psychomotor test performance on temperament assessment in observational situations. J. Nat. Inst. Personnel Res., 1953, 5, 77 - 84.

60. REUNING, H. A new flicker apparatus for measuring individual differences. J. Nat. Inst. Personnel Res., 1954, 6, 44 - 54.

61. ROSS, J. Predicting practical skill in engineering apprentices. Occup. Psychol., 1962, 36, 69 - 74.

62. SEASHORE, R.H. and R.D. ADAMS. The measurement of steadiness : A new apparatus and results in marksmanship. Science, 1933, 78, 235.

63. SEASHORE, R.H. Work and motor performance. In : Stevens, S.S., ed. Handbook of Experimental Psychology, pp.1341 - 1362. New York, Wiley, 1958.

64. SEASHORE, S. The aptitude hypothesis in motor skills. J. Exp. Psychol., 1932, 14, 555 - 561.

65. SKAWRAN, R.F. The Willemse board as a personality test. Psychol. Afr., 1962, 9, 72 - 85.

66. SPAETH, R.A. and G.C. DUNHAM. The correlation between motor control and rifle shooting. Amer. J. Physiol., 1921, 56, 249 - 256.

67. SPEKE, J.H. Journal of the Discovery of the Source of the Nile. London, Dent and Dutton, 1906. (Original publication, 1863).

68. TEEGARDEN, L. Occupational differences in manipulative performance of applicants at a public employment office. J. appl. Psychol., 1943, 27, 416 - 437.

69. THORNDIKE, R.L. Personnel Selection. New York. Wiley, 1949.

70. TIFFIN, J. Industrial Psychology. New York, Prentice-Hall, 1947.

71. TROMP, J. Selection of apprentices for the South African Air Force. Psychol. Afr., 1962, 9, 60 - 66.

72. VERNON, P.E. The structure of practical abilities. Occup. Psychol., 1949, 13, No. 2, 81 - 96.

73. VERNON, P.E. The classification of abilities. Educ.Res., 1960, 2, 184 - 193.

74. VERNON, P.E. The Structure of Human Abilities. London. Methuen, 1950.

75. WALKER, R.Y. and R.D. ADAMS. Motor skills : the validity of serial motor tests for predicting type-writing proficiency. J. gen. Psychol., 1934, 11, 173 - 186.

Final Overall Discussion and Conclusion

1. Animal behaviour, and particularly that of man, can be roughly but conveniently classified into the following five processes : (1) Sensory, concerned with the reception of physical stimulation. (2) Perceptual, in which the sensory data is interpreted. (3) Motor reactions in response to sensory or perceptual processes. (4) Emotional, involving affective or conative sets. (5) Intellectual, which involves a complex amplification of the perceptual.

For both the study of these processes in order to further scientific knowledge, and the application of the knowledge gained to useful practical purposes such as vocational guidance, personnel selection and classification, the improvement of equipment handled by people, and training, apparatus of various kinds can be a great help, and is often indispensable to provide controlled testing conditions and make accurate, objective measurement possible.

2. A very extensive area in which apparatus is pretty well essential, is that of Sensory-motor activities, where the main purpose is usually the observation and measurement of muscular movements. The term "motor" implies movement, but because the movements of animate organisms are never purely mechanical, these activities are preferably referred to as sensory-motor, perceptual-motor, or psychomotor. Directed or voluntary movements that require skill of some sort are generally of wider interest to psychologists, educationists and employers

of personnel, than involuntary movements, although the latter can often be significant in the diagnosis of physiological or emotional states.

3. The technical designs and experimental applications were described of seven more or less complex apparatus tests intended mainly for measuring the sensory-motor skills of prospective air-pilots. All the tests had useful reliability, and in all there were some measures that showed a significant relationship to success in flying training. However, two of the tests that involved very fine patterns of muscular control, namely Handlebars Co-ordination and Steadiness, appeared to have the least usefulness for predicting success in flying. The correlations between them (Time, 0.53; Error, 0.40) were exceptionally high for sensory-motor tests administered to a sophisticated group.

The validity coefficients obtained with these tests were probably depressed by the restriction of range in the criterion groups through pre-selection. This would have applied particularly to Specific Arm-leg Co-ordination, Steadiness, and Hand-foot Reaction, which were included in the actual selection battery.

In Steadiness and Variable Co-ordination there was some evidence that the threat of physical punishment for incompetence, in the form of moderately unpleasant electric shock, was beneficial to the validity of a test in relation to a physically hazardous occupation like flying.

Although the tests in Section I were intended mainly for air-pilot candidates, some of them might have wider uses in both basic research and occupational selection.

4. Of the thirteen sensory-motor tests intended mainly for industrial research and classification, three were "Assembly" tests involving both sensory-motor skill and the ability to perceive three-dimensional form relations. Twelve of the tests showed useful reliability. It was not possible to assess the reliability of the "Three-dimensional" Peg-board, because it was given as one continuous task.

5. The two tests of fine Implement manipulation, namely Chopsticks and Tweezer-nozzle, correlated significantly with a factory operation of "Nutting". Chopsticks showed small significant relationships with Secondary Function, Emotional Stability, Visual Selective Reaction, Visual Perception Rate, Heart and Respiration Rate. A coefficient of -0.36 was obtained between Chopsticks and Pyruvic Acid level after physical work, probably attributable to biochemical factors influencing behaviour. Chopsticks was found useful as a situation for observing temperamental manifestations.

On Tweezer-Nozzle, where it was possible to compare the respective performances of Africans and Europeans, no significant differences were found between the groups. Good performers tended to estimate their achievement more realistically than inferior performers. Most subjects were inclined to over-estimate what they could actually achieve, and the over-estimators tended to be less realistic. Mild to moderate alcoholic intoxication did not affect performance.

6. The six simple skill tests, namely Tweezer-Nuts, Tweezer-Mirror, Two-hand Sticks, Steadiness Stick, Tray Co-ordination and Ring Throwing, provide a wide variety of manipulative patterns, are also suitable for group administration, and have the advantage of being comparatively cheap to reproduce in quantity. These tests showed no validity for selecting operatives in a porcelain factory, which could partly have been due to the lack of specific and dependable criteria. The respective skills tapped by them tended to be highly specific, with a mean intercorrelation of only 0.15. Among the less intelligent subjects there was more overlap between mental and sensory-motor abilities than among the more intelligent.

7. Of the tests administered to mine-workers, Object Sorting and "Three-dimensional" Peg-board discriminated well between the Non-mechanical, Mechanical, and Supervisory categories; and Tripod-and-weight, between the Non-mechanical and Mechanical, but not between the Mechanical and Supervisory. Object Sorting, Tripod-and-weight, Dynamic Coincident Reaction, and Code-signal Reaction, all showed substantial validity for winch-driver selection.

8. The directed muscular activities of man are numerous and diverse. Whereas those of a gross athletic kind have been shown to intercorrelate quite highly, others requiring smaller and more precise movements are, for the most part, very specific. In heterogeneous or low-grade groups relationships between Sensory-motor tests, and between the latter and mental tests, tend to be larger than in selected high-grade groups.

Generally, correlations between Sensory-motor and Mental tests are small and positive, but can be more substantial in the case of complex Sensory-motor tasks. Correlations between speed of work and accuracy in Sensory-motor tasks, are sometimes zero, but usually slightly negative, and more substantial in complex tasks requiring very precise muscular control.

The theory, originally proposed by R.H. Seashore, that relationships between sensory-motor activities depend more on the patterns of movement than on the musculature or sense-modalities employed, is supported by the results of the writer and others. While there appears to be no general Sensory-motor factor, evidence points to the existence of rather narrow group factors, and it is likely that there are broader group factors of "personal tempo", as distinct from actual skill.

9. In the field of Visual-perception, where motor responses are usually of slight or even negligible importance, apparatus can also contribute usefully. Examples are the Locus-estimation Test for selecting ground-control interceptor operators in the Air Force; the Cancellation Timer for administering various printed Cancellation tests; and the Dual-purpose "Fall" Tachistoscope for presenting various visual tasks requiring the exposure of either moving or stationary stimuli for short periods.

10. Performance in the Locus-estimation Test correlated 0.40 ($P=0.07$) with the operational criterion. As the only "practical" test among others that were all of the pencil-and-paper or pictorial kind, it was very specific and made a unique contribution to battery validity.

a validity of 0.48 ($P < 0.01$). There was no significant correlation between Speed of Perception and Span of Attention. The former appears to relate more to physical reaction and co-ordination, and the latter to intellectual activity.

13. Although Flicker-fusion is a phenomenon primarily involving visual-perception, experimental findings suggest that it might have wider psychological significance, particularly in the field of Temperamental measurement. The reliability of tests administered with the Compact Flicker-fusion Machine to air-pilot candidates was very high, although it was probably boosted somewhat through the use of the method of limits. Foveal measurements were more reliable than peripheral. A correlation of 0.51 ($P = 0.003$) was obtained between high CFF and the tendency to free expression of emotion, as rated by independent observers in three other test situations. There appears to be a significant relationship between high CFF and success in flying training. The most valid measurement was the Foveal which gave a mean biserial correlation of 0.50 ($P = 0.00003$) with the Success/Suspension criterion. Candidates having a high CFF are possibly more free from some sort of temperamental inhibition (Secondary-function) that hampers those with low CFF. In the light of Wiersma's findings with colour-fusion, it would seem that the Manic temperament is more desirable than the Melancholic for success in flying training.

14. Even in the realm of higher cognition, which is traditionally associated with printed tests, apparatus can play a useful rôle by providing standard, practical situations

for measuring ability in tasks of substitution or trial-and-error learning. A test of Paired-associate Learning administered to air-pilot candidates by means of the Multiple-choice Instructional Device, correlated with certain branches of flying training, particularly the more practical; the most significant coefficients being -0.43 (Approach aids) and -0.41 (Flying per cent) with "Time"; and 0.47 (Hours to solo) with "Errors". The test showed small to moderate correlations with certain other tests, intellectual, perceptual and sensory-motor, but none with temperament and personality variates of the introversion/extraversion kind. Learning was influenced by both negative transfer through experimental sequence, and differences in the degree of associative affinity between stimulus items and response references. Material that was temporally nearest during the learning process tended to be most intimately associated in the recall task.

The general effect of Perseveration, as measured on the Serial Discrimeter emerged as a tendency for subjects to take longer and make more mistakes when they had to deal with a substitution task involving changes in the normal colour association between stimuli and responses. Error-proneness in this activity correlated 0.46 (P = 0.01) with Suspension in flying training. In the Learning section on the Serial Discrimeter, Error correlated 0.38 (P = 0.05) with Suspension. The correlations between Perseveration and Learning were : Errors, 0.22; Time, 0.55.

In both the Learning tasks administered respectively with the Multiple-Choice Instructional Device and the Serial

Discrimeter, Rate of work showed a wider overlap with other tests, than Error, and is probably a broader personal characteristic of the individual.

Secondary-function as reflected by measures of Flicker-fusion threshold did not correlate significantly with perseveration measures on the Serial Discrimeter.

15. In most of the intellectual tests administered to air-pilot candidates, English-speaking candidates as a group tended to perform better. In two perceptual tests, namely, Repeated Letters and Dots-cancellation, which involve sustained attention of the "clerical" kind, the Afrikaans-speaking were superior. A similar result has been obtained by Dr. Reuning of this Institute with another "clerical" test of sustained attention, the Pauli. To a slight but significant extent the Afrikaans-speaking were also superior in the Perseveration test.

In the tests of Paired-associate learning administered through practical media, and several Sensory-motor tests examined so far, no significant differences between the groups were found. Two possible explanations for the superiority of the English-speaking group in the mental tests are, firstly, there is a far greater quantity and variety of technical and scientific literature available to the English-speaking in their own language; and secondly, military air-pilotage as a career has less attraction for the intellectually higher-grade Afrikaans-speaking than for the equivalent English-speaking.

16. For purposes of occupational selection or classification, practical tests, which usually involve apparatus of some kind,

have both drawbacks and advantages. They are less convenient and economical both to produce and administer than printed tests, and their scope is often restricted because many practical abilities tend to be very specific. Good "general-purpose" tests of the practical sort, particularly those where the emphasis is on manipulative skill, are comparatively scarce. Practical tests which appear to have the most "general-purpose" value are usually hybrids in which much of the variance is attributable to perceptual, cognitive and even temperamental factors. While practical abilities can be most efficiently predicted by specific "job" tests, their use in selecting the industrial rank-and-file is seldom economically justifiable, and the search for instruments having the best possible "general-purpose" value is likely to continue, even if it often proves to be something of a wild-goose chase.

As a rule the more successful and widely used apparatus tests are constructionally simple. The use of elaborate equipment is only worthwhile for high-level or vital selection, and even here, complication beyond a certain point is undesirable.

The principal justification for the use of practical selection tests, is that there are many practical factors which cannot be sampled by pencil-and-paper or pictorial tests. Even a small degree of validity can warrant the application of a test if it makes a unique contribution to a selection battery. Ultimately, however, the value of a test depends on overall expediency.

As orientation towards concrete reality is the greatest common factor pervading all cultures, practical tests are usually less "culturally inappropriate" for African subjects than pencil-and-paper or pictorial tests. Moreover, they can be demonstrated to polyglot groups with the minimal need for verbal explanation, and they also enable the subjects to perceive a more obvious connection between test and job, which appears to favour motivation.