THE UNIFICATION OF HARDNESS MEASUREMENT

1991

BUREAU INTERNATIONAL DE METROLOGIE LEGALE
11, rue Turgot - 75009 PARIS - FRANCE

PREFACE

Research results in the field of the metrology of hardness published in the last decades in technical journals and conference reports in various languages have been collated in three publications of BIML:

Factors influencing hardness measurement, 1983
Hardness test blocks and indenters, 1984
Hardness standard equipment, 1989.

The present publication complements the earlier ones, dealing with subjects of rather basic character, such as

- hardness as a special measured quantity,
- ways of assuring metrological control,
- methods of uncertainty evaluation,
- international hardness comparisons.

The aim of previous publications has been maintained: to survey the dispersed research results from this narrow field of metrology with a more or less unified presentation. When necessary the reader may study the original publication for more detail. The series is intended to be of special help to metrologists who are about to start work on some detailed problem of hardness measurement , specially standardizing work, to give them a start in the search of the literature.

The author wishes to express his appreciation to the research workers listed in the Index, whose results are summarized. Any comments, proposals or additions to the text or to the references will be welcomed and appreciated.

Dr Ferenc PETIK
Assistant Director

CONTENTS

1. The special nature of hardness measurement

A hardness value is the result of an experiment performed under standardized conditions and based on a convention.

1.1. The place of hardness measurement in metrology

The inclusion of hardness measurement in the broad system of measured quantities is often a subject of dispute, of objections of principle, and of other difficulties among metrologists. The material-testing specialist who measures various quantities tries to find interconnections, relationships and similarities between the measurement of hardness and that of other quantities. To clarify these problems, it is useful to begin with the examination of some of the principles of hardness measurement, perhaps from a somewhat philosophical approach.

Hardness measurement has a special place in metrology. It is useful to look for analogies with some SI-units, to examine the problems of establishing measurement scales, the role and the characteristics of measurement standards.

The process of hardness testing is an experiment performed under prescribed conditions whereby the measurement is performed, after the indentation process, in two steps:

- The primary measurement is length measurement (indentation depth, length of diagonal, diameter). The obtained length value is an auxiliary quantity of hardness measurement.
- The second step of hardness measurement is the determination of the hardness value from the auxiliary quantity by employing the equation defining the hardness testing method.

The above mentioned objections and disputes concern always the second step of measurement, and never the primary length measurement which serves as the basis.

The indirect measurement method of hardness is analogous to viscosity measurement: Viscosity is determined in an experiment where, primarily, the duration of liquid flow through the instrument is measured. The second step is the determination of viscosity value from the measured time.

A precondition for measurements of this type is an unequivocal correlation between the measured quantity and the auxiliary quantity determined in the experiment during the primary measurement.

1.1.1. Hardness as a measurable quantity

A measurable quantity represents some characteristic of a phenomenon or of an object, which can be qualitatively described and quantitatively determined. Hardness is a characteristic of the material which is of great interest in technology.

The <u>definition of a quantity</u> to be measured is not always simple. Length and time are obvious concepts for everybody, but mass is no longer. We learn at school that mass means an amount of material. If we try to define force, the difficulties are still greater.

The concept of hardness corresponds to everyday word usage. Even nonprofessionals have an idea of a hardness scale for everyday needs, from the very soft to the very hard. This idea is based on the characteristic of material to resist the penetration of another body. For technological purposes only those hardness ranges are of interest which are called "very hard" by nonprofessionals. In these ranges difference in hardness cannot be determined without testing equipment.

Measured quantities also can be classified according to whether we have a sensing organ for them, or not. The first group includes e.g. length, force, luminous intensity, the second one the electrical and several chemical quantities, and many more. Technological hardness has a special status also in this respect. As shown above, we have sensing organs for hardness, which are however not of use in the hardness ranges of interest in technology.

The measured quantity may be constant (e.g. a given mass, or the length of a given object), or the quantity may be variable with time (e.g. the length of an object being warmed up). There are measuring instruments capable of following and indicating continuously the continuous variation of the quantity to be measured (thermometer, dial gauge). Measurement methods based on comparison, however, are suited only to the determination of a constant quantity. Hardness of a material, the result of a technological process (mostly heat treatment), can be measured only after the treatment is complete. Hardness is measured in the stabilized state. Testing apparatus is not suited to measuring continuously changing hardness.

On account of the small local destruction of the surface of the specimen, hardness testing is a not repeatable measurement, in common with all destructive methods of material testing. (But even the measurement of an SI base unit, time, is not repeatable). On account of the imperfectly uniform hardness of a given material surface, a sample is taken from the local hardness values on the surface. The repetition of sampling is of course burdened with uncertainty. Also in the case of changing hardness, in consequence of ageing or heat effects, a new sample has to be taken representing the new hardness of the surface.

1.1.2. Hardness standards

In the realization of different units of measurement various methods are employed. For the SI base units and for some derived units the methods can be classified as follows:

a. Material measure: The only unit having a material measure standard is mass, kg. Before 1960 the metre too belonged to this class.

b. Experiment to realize the unit: For time, s; electric current, A; and various derived units, such as force, N, and pressure, Pa, etc. From 1960 to 1983 the standard of metre also belonged to this class (definition by the krypton radiation).

c. Definition by natural constants: For length, m (definition based on the velocity of light); for thermodynamic temperature, K, and amount of substance, mol.

The realization of a measurement standard for hardness would be possible in two ways:

- Definition of the hardness scale by fixed points, realized by standard hardness blocks.

- Definition of the hardness scale by an experiment performed on a hardness standard equipment.

The first hardness scales, e.g. Mohs, were based on fixed points. Even for technological hardness scales used in our day various similar solutions have been proposed. However, as is well known, earlier discussions of this question were resolved in favour of the second solution which corresponds to method b. above and is so in some way related to the standards of the units s, A, N and Pa.

The blocks are not called standard hardness blocks any more, but standardized (or calibrated) hardness test blocks. The blocks are nevertheless indispensable as secondary transfer standards for hardness values, in the traceability of measurements performed by immobile equipment.

1.1.3. SI-scales, reference-value scales

Measurable quantities are either physical quantities given in SI units, or quantities given on a reference-value scale. According to the International Vocabulary of General Terms in Metrology a reference-value scale is, "for a given quantity or property a series of values determined in a defined manner and adopted by convention". Hardness scales belong to this group, together with numerous other quantities that are very important for science and technology, e.g. the practical temperature scale, the time scale (the time that elapsed since a given datum), viscosity, the International Sugar Scale, various scales of alcohol content, light sensitivity of films, wind force scale, seismic intensity, and many others. These scales are indispensable for the knowledge of the world, and so cannot be considered as part of "second-order metrology".

Quantities given on a reference-value scale are:

- on the one hand, sufficiently well known to express relationships by numerical values,

- on the other hand, not sufficiently well known to describe and generalize the relationships with the exactness of physical laws.

With the progress of science a quantity given on a reference-value scale may later become measurable on an SI-scale, i.e. may be transformed to a physical quantity. E.g. the Celsius and Fahrenheit temperature scales are reference-value scales. Thermodynamic temperature (Kelvin scale) in turn, is a physical quantity, its unit belongs to SI. Or one of the former units for kinematic viscosity was the Engler degree; now we have the m^2/s in SI.

In hardness measurement too, the first timid signs of a development in this direction have appeared: It is not correct to specify Brinell hardness with the symbols kgf/mm² or N/mm², like the physical quantity pressure. There are two reasons for this: The distribution of pressure under the ball indenter is not uniform. Furthermore, the area which is taken into consideration for calculating the hardness value is measured only after the test load has been removed (recovered indentation). We have no knowledge of the magnitude of the indentation area under load when performing a Brinell test. With hardness testing under load by the Vickers method, which is often mentioned lately, it is possible to correlate hardness with mean pressure under the indenter, since the test load and the indentation area produced by it exist simultaneously. Thus this is a small step in approximating hardness measurement to SI.

1.1.4. Some characteristics of hardness scales

The following characteristics of physical quantities can be utilized in application:

- Relation (two measured quantities are equal, or one is larger than the other).
- Addition (by adding two measured quantities of the same kind, we obtain a new quantity of the same kind).
- Multiplication (by multiplying a measured quantity by a positive number we obtain a new quantity of the same kind).

Examples of employing the three characteristics for an SI-unit:

$$3\ g = 3\ g,\ 3\ g < 4\ g$$
$$3\ g + 4\ g = 7\ g$$
$$5 \times 3\ g = 15\ g$$

For reference-value scales, only the first characteristic applies, e.g.

650 HV = 650 HV and 650 HB > (harder than) 550 HV.

It is, however meaningless to try to perform the following operations:

250 HV + 200 HV = ? (Increase the hardness by 200 HV?)
3 x 250 HV = ? (To make three times as hard?)

Hardness scales are not linear, accordingly even substraction, the operation with scale intervals is impossible:

$$450\ HV - 440\ HV \neq 460\ HV - 450\ HV \neq 10\ HV.$$

The impossibility of multiplication of hardness scales has another consequence: For physical quantities,

Value of the quantity = numerical value x unit
5 m = 5 x 1m

The numerical value indicates, how many units compose the quantity. In the case of reference-value scales this principle is not applicable:

$$55 \text{ HRC} \neq 55 \times 1 \text{ HRC},$$

since HRC is not a unit of measurement, but the symbol of the conventional measurement method.

Hardness scales have no real zero. On the Rockwell scale 0 HRC is in principle feasible, namely when the permanent increase of depth of indentation is 0.200 mm. (Even negative HRC value can be imagined according to the definition of the method). However this zero is quite arbitrary and useless in practice. It is prescribed namely that the Rockwell-C method is applicable only in the range 20-70 HRC.

The zero of the Vickers hardness scale would correspond, according to the definition of the method, to an infinitely large indentation, that is that the indentation process never ends, is not stabilized. Technically this is also meaningless.

The limited possibility of conversion of values of different hardness scales follows from the precedings. The relationship between different scales for physical quantities is linear. E.g. length can be measured in metres or feet. The relationship between the two length scales is given by the factor 3.25. The relationship between different hardness scales is not linear. A remarkable proof of this fact is supplied by diagrams indicating the uncertainty of hardness standard equipment as a function of hardness levels [P-28]. Rockwell hardness uncertainty decreases as hardness increases (± 0.2 HRC at 30 HRC and ± 0.1 HRC at 65 HRC). For Vickers hardness the relationship is inverted (± 1 HV at 150 HV and ± 3 HV at 600 HV). This is, however, no physical or metrological contradiction. The reason is in the different character of the two scales. Scale intervals are smaller on one scale in that hardness range where the intervals of the other scale are larger. The different definition of various hardness testing methods is the reason why the conversion of hardness values to another scale can be performed only with the help of experimentally produced tables, and the conversion formula established for a certain material is not necessarily valid for other materials.

1.2. The metrology of hardness testing

The most important subjects of research work in the metrology of hardness can be summarised as follows:

- Establishing, maintaining and comparing hardness reference scales,
- Factors influencing measured hardness values (instrumental and personal),
- Errors at various levels of the hierarchical order of hardness measurement.

1.3. A short summary of the history of research work in the metrology of hardness

The Brinell method appeared in 1900 and remained the only hardness testing method available for 20 years. The need to test heat-treated materials led to the development of the Vickers method (in its present form in 1925). But in the meantime the need to simplify hardness testing and make it less time consuming than the methods employing microscopes led, in 1920, to the Rockwell method, with the straightforward dial-gauge reading.

At first, hardness testers were produced by a single, and then by a few manufacturers who were careful to maintain the consistency of measurement of the testers they sold, and users had confidence in the apparatus as supplied by the manufacturer. As the number of manufacturers increased, the first step towards unification was the development of standard specifications for methods of hardness testing in the years between the two World Wars. These standards contained relatively few data on test equipment.

When disputes arose about differing values of hardness measured on the same object at different places, hardness-testing specialists felt the need for the metrological assurance of their test results, though this expression was not yet in use that time.

The obvious solution was to procure sufficiently uniform and stable steel blocks which could be used for comparing the results of measurement of different hardness testers. Hardness test blocks were available in 1937 in Sweden and Germany, the literature refers to similar blocks in Japan two years later [F-3, W-13, Y-1].

The blocks were well suited to eliminating some disputes, but where there were discrepancies between the results of two hardness testers the question remained: which tester gave the correct result? The traditional metrological solution was evident; a standard measuring instrument which had superior metrological characteristics to those of the commercially available hardness testers.

The first standardizing machine was built by K. Meyer [M-11] and came into operation in 1943. In the post-war years, efforts aimed at the unification of hardness measurement gathered momentum; more and more countries and metrological institutions started work in the field. Hardness testing was empirical and piecemeal originally, but later systematic research was started.

A schematic drawing of a Brinell standardizing machine was published in 1949 in the Soviet Union (reproduced in [P-29]). Philips and Fenner [P-13] published a detailed description of the Rockwell standardizing equipment of the NPL in England in 1951. By 1955, standardizing machines with deadweight loading (mostly Rockwell) were in operation also in both German states, Austria and Japan. In the USA and Sweden, commercial testers of improved quality were used as standardizing machines. Officially calibrated blocks were available in Germany from 1952. The use of these blocks in a national survey of the state of industrial hardness measurement practice revealed a very sad situation [M-11], but this national comparison gave impetus to the development and use of standardized blocks, and of checking equipment. For example, a description of measuring equipment and a method for checking the geometry and surface quality of diamond indenters was published in 1951 by Tolmon and Wood [T-1].

Early research work

The effects of indenter geometry were analysed theoretically and empirically by Hild, Guidel and K. Yamamoto [H-4, G-1, Y-8]. Extensive research on indenters and on all factors influencing hardness results, including the quality of blocks, was performed in the years 1951-1956 in Japan. It is most regrettable that only very condensed summaries on this work are available in languages other than Japanese [Y-13, Y-14].

Standardizing machines were equipped with selected indenters, and correction curves were established for each of them. Hormuth [H-9] introduced a group standard of indenters to realize the hardness reference scale. The first international hardness comparison, made in the years 1954 to 1957, compared the standardizing machines of Austria, France, GDR, FR Germany, Japan, Sweden, UK, USA and the USSR [M-11]. Discrepancies which were revealed gave rise to a proposed international convention on the unification of hardness values as early as the late 1950s.

In addition to the Rockwell standards, more and more Vickers and Brinell standardizing machines came into use. It was soon recognized that the weak point of these machines was the measuring microscope. In various countries, work teams were set up to perfect the optical and metrological characteristics of the microscopes used for the measurement of indentations.

The analysis of errors in hardness measurements was also started (Piliptchuk [P-14]). As the hardness test is a measurement which is not repeatable at the same place, but is easily repeatable at another point on the same specimen, the possibility of exploiting the rapidly developing statistical methods of planning experiments and analising results was soon recognized (Marriner, [M-1]).

(To avoid undue omissions and unintentional offence in connection with this brief historical survey of hardness research work, only those research workers who are no longer active are mentioned once by name. But even the list of countries in which work has been done cannot be exhaustive. The development of science and technology is always the result of numerous contributions, frequently simultaneous and in parallel. The author apologizes for any omissions.)

1960-1970

In the 1960s, Japanese research workers introduced statistical analysis for every aspect of hardness standardization. The international comparison which started in 1961 was planned from the beginning in such a way that additional information might be made available by the application of mathematical statistics [Y-6]. This new tool of metrologists was soon complemented by another, the computer.

This hardness comparison was not the only one made in the 1960s. Various groups of countries, regional organizations (e.g. CMEA) organized comparative measurements between standardizing machines with and without indenters, between indenters, between measuring microscopes, etc. [P-4].

Various efforts were made to improve the metrological characteristics of the spiral microscopes generally used in Rockwell standardizing machines (e.g. in FRG, Czechoslovakia, GDR, Hungary), but the classical method offered no real solution. Only technological development (the laser) permitted a solution to be found in a different way, but that came some years later.

All the research themes started in 1950s were continued, with an enlarged number of participating institutions in a greater number of countries. In Japan the new Rockwell standardizing machine employed a lever transmission for loading. Around 1965, much attention was given to the parameters of the loading cycle (e.g, in Romania, FRG, Japan, UK). The surface quality of diamond indenters was the subject of some research work (Poland, GDR, FRG). Crystallographic characteristics of the diamond were established and taken into consideration during the manufacture of indenters. The functional examination of indenters permitted useful analyses (among others, Hungary and Poland).

In standardizing Vickers machines efforts were made to obtain a standard indentation of specified diagonal length (Germany, Czechoslovakia, Japan). Shore-hardness standardizing machines were set up in Japan, where also the influence factors of the Superficial Rockwell test were examined. Much research work was done in various countries in connection with the interconversion of hardness values established by different methods. The introduction of SI, of the new unit for force, the newton, also caused some discussion in hardness-testing circles, but this question was peacefully solved; animosity has long since disappeared.

The systematic analysis of sources of error in standardizing machines was the subject of various studies (in Japan, Czechoslovakia and Hungary). The state of industrial hardness measurement practice was the subject of new analyses in various countries (Czechoslovakia, Japan, FRG, GDR).

<u>1970-1980</u>

In the following decade, several existing Rockwell standardizing machines were equipped with laser interferometers to measure the indentation depth. On newly installed standardizing machines the design was improved by the introduction of frictionless bearings and other newly developed machine elements (e.g. in Italy and USA). The standardized hardness test block became a mass product, and the questions of the costs of production and calibration gained importance. At the same time the choice of commercially available high-quality blocks increased.

Several countries (Soviet Union, Czechoslovakia, GDR) published standard specifications for the national hierarchy scheme of hardness measurements. The unified way of stating the uncertainties of hardness measurements made by the standard equipment, transferred by the blocks, and measured by the ordinary measuring instruments (commercial hardness testers) was also studied (Hungary). In connection with Vickers standardizing work and microhardness testing, much attention was given to personal and physiological effects when using measuring microscopes (Japan, Italy).

In the field of industrial hardness testing, the requirements of automated production lines became of great importance. The increasing role of plastics in technology and everyday life also demanded new departures in hardness testing.

Present

As we reach the present time in this historical survey, only two developments are mentioned here. One is the endeavour to eliminate the human element from the measurement of Vickers indentations. The other is the great number of proposed new hardness-testing procedures (e.g, series measurement, test under load, etc.). New methods used to be developed with the aim of eliminating some of the drawbacks of the classical methods. Generally, a new method does not easily take root. A good idea needs a manufacturer ready to bear the costs of developing and marketing the new tester. It is a long road to obtaining general acceptance, which can only be as a result of perceived advantages over earlier methods. The acceptance of a new method becomes complete when standardization institutes are ready to write a standard specification for it.

It is worth mentioning that most of the new methods proposed in recent decades have been indentation methods. The great hopes of earlier times concerning non-destructive hardness testing (ultrasound, magnetic, etc.) seem not yet to have been realized.

* *
*

The results of this research work were summarized in the surveys contained in four BIML publications [P-21, P-22; P-28] including the present one.

The aim of all this work is the metrological assurance of hardness testing.

2. The uniformity of hardness measurements

2.1. Ways of assuring metrological control of hardness measurement

The elements of assuring metrological control of hardness measurement are the following:

a) Measurement units, scales

- To maintain national or international hardness measurement scales, hardness standard equipment (standardizing machines, including measuring microscopes) is installed and maintained under constant metrological control, to ensure stability of hardness scales and its reproduction with the best possible uncertainty [P-17, P-28].

b) Transfer standards

- To ensure the connection between stationary standard machines and stationary hardness testers, hardness test blocks of specified quality are produced and calibrated (standardized, verified) on the standardizing machines [P-22].

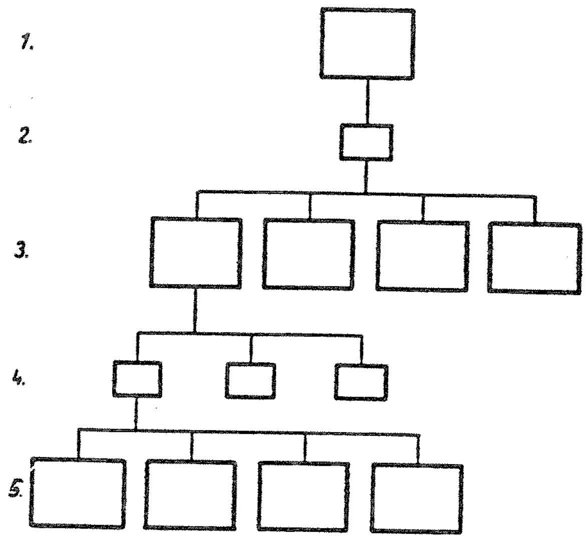

Fig. 1. Hierarchy scheme of hardness measurement, after the setting-up of an international standard. 1) International hardness-standard equipment. 2) and 4) Standardized blocks. 3) National hardness standard equipments of different countries. 5) Hardness testing machines in industry (working instruments).

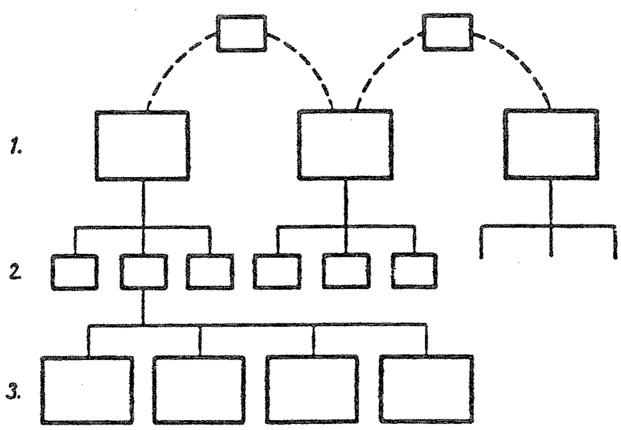

Fig.2. Hierarchy scheme of hardness measurement. Present state. 1) National hardness-standard equipments of different countries. 2) Standardized blocks. 3) Hardness testing machines in industry (working instruments).

c) Industrial hardness testers

These measuring instruments (testing machines) are subject in conformity with national legal prescriptions, to

- pattern examination and approval,
- initial and periodic verification (calibration).

The most important steps of the verification process are:

- Examination of the installation of the machine, general check of the state and functioning of the machine.
- Verification of the test force.
- Verification of the indenter (geometrical control and functional examination by measurements on a standard machine).
- Verificatioon of the measuring device.
- Indirect verification of the hardness tester by means of standardized hardness test blocks.

All these operations should be based on metrological specifications, preferably internationally harmonized prescriptions (ISO International Standards, OIML International Recommendations).

2.2. Hierarchy schemes

The values determined by all measuring instruments, thus also by hardness testing machines, must be traceable to the standards. The connection between the international and national primary standards, and the instruments used in everyday measurement practice (working or ordinary instruments) is ensured by the metrological institutions by the methods indicated in hierarchy schemes. The scheme is a useful means of indicating the various classes of accuracy. Instruments standing in a lower row have a lower uncertainty and accuracy than those standing higher, because each instrument is checked by the instrument standing in the immediately higher row. The principles for the establishement of hierarchy schemes are discussed in detail in OIML International Document D5 (Edition 1982).

The desirable future hierarchy scheme for hardness measurement is shown in Fig. 1. This is similar to the scheme of physical quantities which have an international standard. In Fig. 1 the large squares indicate the stationary equipment, the small squares the portable blocks, which ensure the connection between the stationary machines.

Though the necessity of setting up international hardness standards has been discussed for decades, these efforts have been without success until now. (Only some state groups, e.g. CMEA, have regional international standards). At present in most countries the national hardness standard represents the highest element in national hierarchy schemes.

The present state of the hierarchical order is shown in Fig. 2. The national hardness-standard equipments stand in row 1. These are maintained at present independently of each other, the only connection between them being the international comparisons performed with the help of standardized blocks. On the national hardness-standard equipments

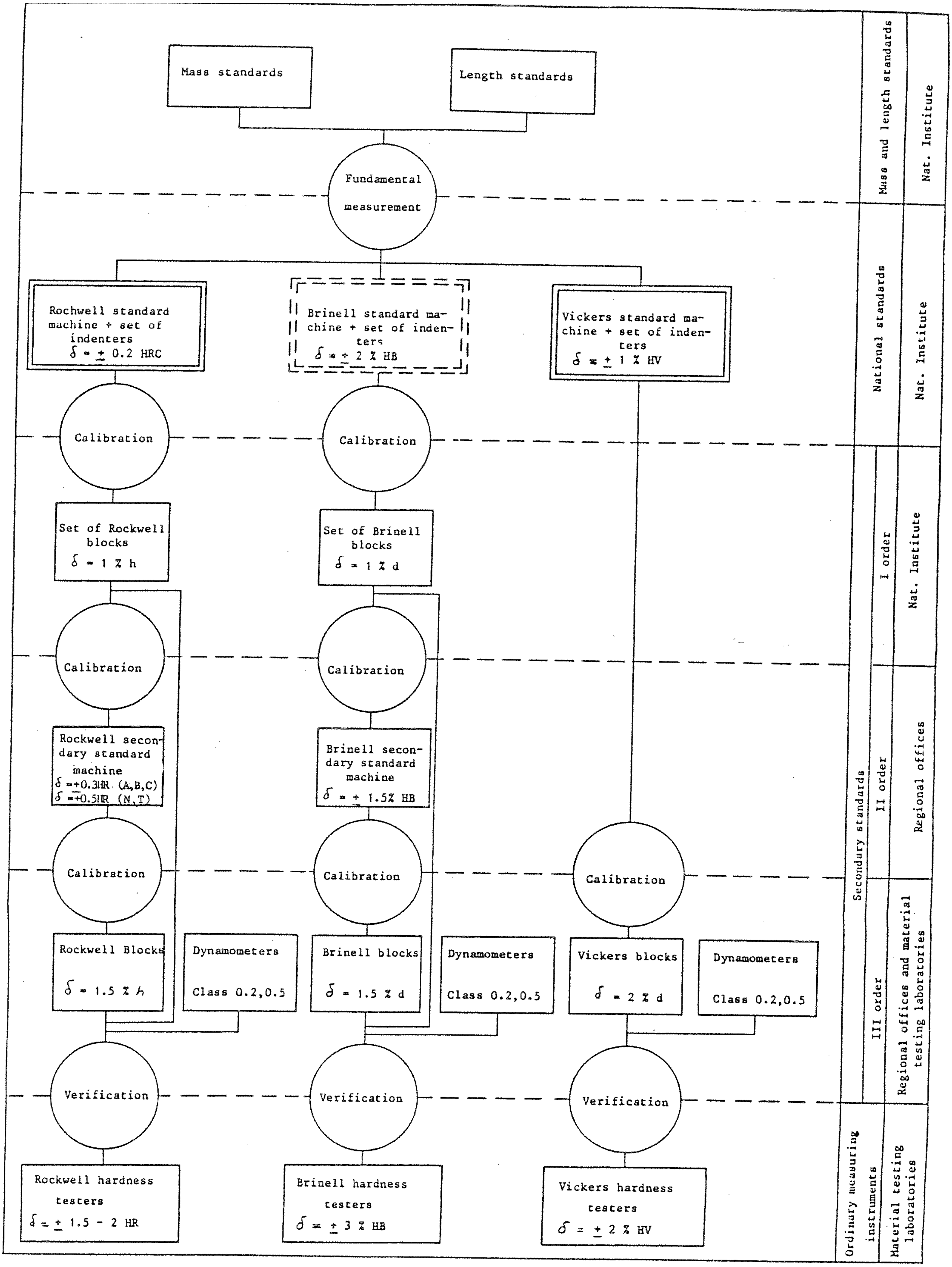
Mass standards
Length standards
Fundamental measurement
Rochwell standard machine + set of indenters δ = ± 0.2 HRC
Brinell standard machine + set of indenters δ = ± 2 % HB
Vickers standard machine + set of indenters δ = ± 1 % HV
Calibration
Calibration
Set of Rockwell blocks δ = 1 % h
Set of Brinell blocks δ = 1 % d
Calibration
Calibration
Rockwell secondary standard machine δ = +0.3HR (A,B,C) δ = +0.5HR (N,T)
Brinell secondary standard machine δ = ± 1.5% HB
Calibration
Calibration
Calibration
Rockwell Blocks δ = 1.5 % h
Dynamometers Class 0.2,0.5
Brinell blocks δ = 1.5 % d
Dynamometers Class 0.2,0.5
Vickers blocks δ = 2 % d
Dynamometers Class 0.2,0.5
Verification
Verification
Verification
Rockwell hardness testers δ = ± 1.5 - 2 HR
Brinell hardness testers δ = ± 3 % HB
Vickers hardness testers δ = ± 2 % HV
Mass and length standards
Nat. Institute
National standards
Nat. Institute
I order
Nat. Institute
Secondary standards
II order
Regional offices
III order
Regional offices and material testing laboratories
Ordinary measuring instruments
Material testing laboratories

Fig. 3

blocks 2 are calibrated which are employed in turn, for checking hardness testing machines used in industry designated by 3.

The principles given in these figures were elaborated more in detail in some hierarchy schemes used by national metrology services or in regional organizations.

The hierarchy scheme elaborated in Poland [M-26] is shown in Fig. 3. The national hardness standards are traceable, by fundamental measurements, to the national mass and length standards. The Brinell hardness standard machine is not existing yet. The set of secondary standard Brinell blocks of the I order are calibrated abroad. The scheme indicates that Rockwell hardness testers in the industrial laboratory can be verified either by blocks of the III order, calibrated on the secondary standard machine of a regional metrological office, or by blocks of the I order, calibrated on the national standard machine.

In a small country (e.g. Hungary) it is possible to calibrate (standardize) all blocks necessary for the verification of hardness testers on the standard machine. In terms of Fig. 3 this means that the rows of secondary standards of the I and II orders are not necessary. This can be called a three element scheme (see Fig. 2). In the case of Poland and of larger countries secondary standard machine(s) are inevitable, the scheme includes five elements in the hierarchy line. SMOLITCH [S-8] showed on practical examples that the longer traceability line reduces the reserve of uncertainty which remains for the industrial hardness testers.

A principle employed for national hardness standards in several countries is the unity of three elements ("tri-lateral standard") [C-13]:

Machine

Indenter —— Set of blocks

If any one of the three elements is damaged or deregulated (e.g. machine failure, broken indenter, lost block, or block of bad constancy), the standard reference value can be reconstituted by means of the remaining two elements.

This principle is employed in the hierarchy scheme for Rockwell measurement TGL 31543/31 elaborated by the Institute ASMW, East Berlin, the practical realization of which is shown in Fig. 4 [H-14]. The national standard consists of the standard machine and of five standard indenters, level (1) in Fig. 4. The uncertainty of transmitting the hardness value from the standard machine to blocks being calibrated is characterized by the standard deviation s_H and by the unknown error component Θ_H of the standard.

$$s_H \leq 0.05 \text{ HR (A,B,C)}$$
$$\Theta_H \leq 0.15 \text{ HRA}$$
$$\leq 0.25 \text{ HRB}$$
$$\leq 0.15 \text{ HRC}$$

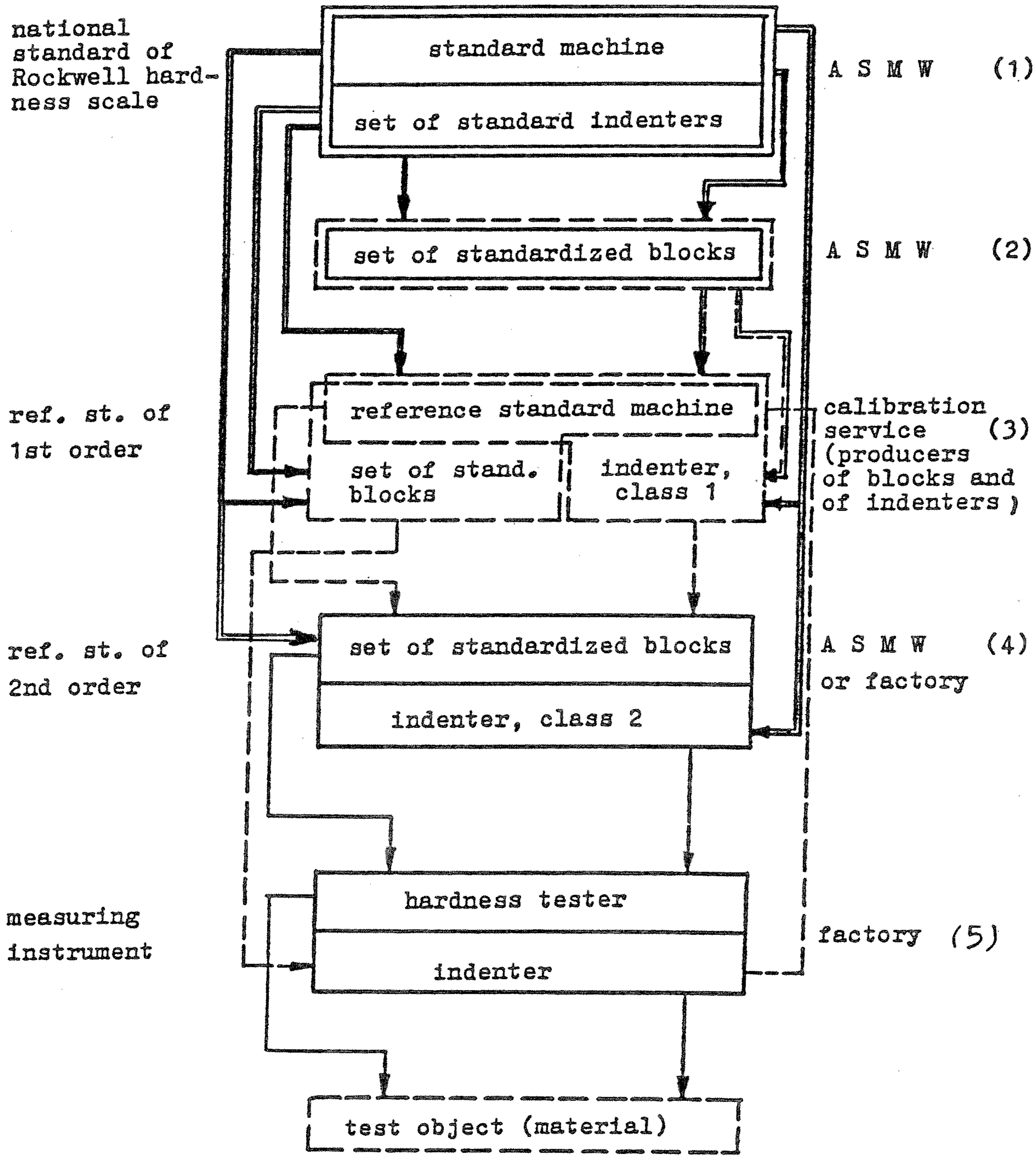

Fig. 4. The practical realization of the ASMW hierarchy scheme according to the standard specification TGL 31543/31.

The set of blocks shown at level (2) serves as the third element of the tri-lateral standard.

The I order reference standard is conserved at regional calibration services and manufacturers of blocks and indenters. This reference standard, shown at level (3), consists again of the above-mentioned three elements. The machine is compared with the national standard by means of the indenters from level (1) and the blocks from level (2).

The indenters class 1 and the set of blocks which belong to the reference standard, are calibrated on the national standard.

The set of blocks consists of 5 elements for HRC, and of 3 elements for HRA, or HRB measurements. These blocks are of better quality than those used for current calibration (permitted range of values approximately halved). The correction value of reference indenters, class 1, should be $\varphi_H \leq 0.3$ HRA or HRC.

The II order reference standards, level (4), consist of sets of blocks and indenters. These are kept by calibration offices of ASMW, or by the quality control organs of factories producing hardness testers. These blocks and class 2 indenters are previously calibrated either at level (3), or (1). The reference indenters, class 2, may have a correction value $\varphi_H = 0.5$ HRA or HRC at the maximum. The set of blocks should contain at least 3 pieces for HRC and 2 pieces for HRA or HRB.

Hardness testers newly produced in factories (initial verification or calibration) or already in use in material testing laboratories (periodic verification or calibration) which are shown as level (5) in Fig. 4 are checked by the blocks and the indenter from level (4).

The hierarchy schemes for hardness measuring instruments elaborated by the CMEA (NTM SEV.41-89) is an example of the calibration chain based on an international (regional) standard. The elements of these schemes are enumerated in Table 1.

Table 1

CMEA hierarchy schemes (for HB, HV, HR-A, B, C, HR-N, T)

1.	CMEA standard (machine, with measuring equipment and indenter)
1a	Measurement
2	National standard (set of blocks + Standard equipment serving as a comparator)
2a	Comparison on the comparator
3	Secondary standards (set of standardized blocks + standard indenter)
3a	Comparison on a comparator
4	Working standards (set of blocks)
4a	Measurement
5	Ordinary hardness testers

In this scheme only the international standard (level 1) is a standardizing machine. The main elements of standards at levels 2, 3 and 4 are sets of blocks. (In small states level 3 is omitted, if a single standardizing machine is sufficient to calibrate all the blocks used for checking the industrial hardness testers in that country). The national standardizing machine is called in this scheme a comparator, forming part of the national standard (2).

This question needs some clarification.

If we want to compare two blocks, then we make 5 or 10 indentations on each on the same standardizing machine in a short period of time. The comparison mentioned in 2a is not exactly the same. On the regional international standard (1) located in Prague, Czechoslovakia a set of blocks is calibrated (1a) which will be sent e.g. to Budapest to serve as national standard of Hungary (2). If the Hungarian Institute wants to calibrate (2a) a number of blocks (3 or 4), before doing so, to check the correct functioning of their standardizing machine, first 5 or 10 measurements are made on the national standard set of blocks (2). Consequently the comparison indicated under 2a is not an actual comparison of blocks 2 and 3, but rather a simplified notation of two steps: firstly the standardizing machine is calibrated by the help of the national standard blocks (2), secondly on the standardizing machine the secondary standard blocks (3) are calibrated.

In any case, the alternance of stationary equipments and portable blocks, as shown in Figures 1 and 2 is inevitable in any hierarchy scheme for hardness measurement. The denomination of the individual elements in the chain depends on local metrological tradition or legal requirements.

In the complete scheme, the secondary standard (3) consists of a set of standardized blocks with marked hardness value, and of a standard indenter. When the secondary standard (3) and working standards (4) are compared (i.e. the hardness value is transferred to the working standard), a hardness tester is used as comparator (3a) with the indenter forming part of the secondary standard (3). What was said above on the comparison 2a, applies also for comparison 3a.

The nature of "comparison" and "comparator" employed in this hierarchy scheme can be put in proper light by taking an analogy, that of calibrating weights on a self-indicating weighing instrument. [The corresponding concepts in hardness measurement figure in brackets.] Weights of e.g. 1.0 - 1.2 kg [secondary standard blocks] to be calibrated are placed on the weighing instrument [hardness standardizing equipment]. The mass [hardness] value read on the scale of the instrument is marked on the weights [blocks]. The pieces calibrated one after another do not need to have the same mass [hardness] value. To ensure that the calibrating instrument is functioning correctly, in certain time intervals, e.g. once in a week, a mass standard of 1 kg [national standard block of known hardness] is measured on the instrument. Can we call this process a "comparison" of the 1 kg mass standard and of the different weights of the range 1.0-1.2 kg? Perhaps it is better to speak of tracing the mass values to the standard, but the weighing instrument also forms a level of the hierarchy scheme, as follows:

Mass standard - (weighing) - Weighing instrument - (weighing)- Weights to be calibrated.

Table 2

Uncertainties at the individual levels of the CMEA Hierarchy Scheme for Hardness Measurement

Level	Form of uncertainty specification	Brinell	Vickers (HV 1-HV 100)	Rockwell	
				A, B, C	N, T
1	Standard deviation, s_r	0.10-0.13% HB	0.2% HV	0.025 HR	0.05 HR
	Limit of not corrected systematic errors, Θ	± 0.2-0.3% HB	± 0.64% HV	± 0.3 HR	± 0.11-0.22 HR
2	Standard deviation of the comparison of the CMEA and national standards, s	0.13-0.18% HB	0.15-0.35% HV	0.03-0.05 HR	0.06 HR
3	Error limits, at 0.99 confidence level, δ	± 0.5-0.6% HB	± 0.6-1.2% HV	± 0.2 HR	± 0.3 HR
4	Error limits, at 0.99 confidence level, δ	± 0.8-1.0% HB	± 0.8-1.5% HV	± 0.3 HR	± 0.5 HR
5	Maximum permitted error, Δ	± 2-4% HB	± 2-6% HV	± 1.5-2.0 HRA ± 2.0-4.0 HRB ± 1.5 HRC	± 1.5 HRN ± 2.5 HRT

Perhaps this mass analogy helps to understand better the operations included in the hierarchy scheme of hardness measurement (Table 1).

Uncertainties specified for each level of the CMEA hierarchy scheme are given in Table 2. This hierarchy scheme was approved not as an International Standard but as a document on a "Scientific-Technological Method". It was apparently constructed with the aim of uniting and harmonizing several already existing national schemes and practice. In some elements and data of the scheme the effects of inevitable compromise solutions are visible.

A national scheme, elaborated on similar principles as the CMEA-scheme, is that of the Soviet Union, published as state standards GOST 8062-85, GOST 8063-79, GOST 8064-79 [K-18, B-15]. The elements are shown in Table 3. These are similar to those given in Table 1, of course without the elements related to the international standard. Therefore the numbering of levels is identical in Tables 1 and 3. Uncertainties specified for each level are given in Table 4.

Table 3

GOST hierarchy schemes (Soviet Union)

2.	State standard (Standard machine with measuring equipment and indenter).
2a	Measurement.
3.	Secondary standards (set of blocks).
3a	Comparison on a comparator.
4	Working standards (set of blocks).
4a	Measurement.
5	Ordinary hardness testers.

In the United States, in the ACCO-Wilson Standardizing Laboratory the calibration of blocks is made by lever type Rockwell testers. For checking the correctness of these, a dead-weight machine was built [D-2). Control charts of the dead-weight machine and of the lever-type machines were compared every other day a first. After confidence was built up, comparison was changed to a weekly basis. Control charts for the two types of machines over a period of approximately five months show a good similarity. Therefore the dead-weight machine is not used any more for calibrating blocks, it provides an independent control as a duplicate (reserve) standard.

2.3. Propagation of uncertainties in the hierarchical order

The peculiarities of hardness measurement discussed in Chapter 1.1 have an influence also on the propagation of errors, on the accumulation of uncertainties at the different levels of the hierarchy scheme.

Stated uncertainty values for hardness standards were discussed in detail in the BIML Publication - Hardness Standard Equipment, 1989 [P-28]. In comparison with standards of most SI-units the uncertainty of hardness standards is relatively high.

Table 4

Uncertainties at the individual levels of the GOST hierarchy scheme for Hardness Measurement

Level	Form of uncertainty specification	Brinell	Vickers (HV 1-HV 100)	Rockwell A, B, C	Rockwell N, T
2	Standard deviation, s_o (n = 10) Limit of not corrected systematic errors, Θ	0.1% HB ± 0.3% HB	0.1-0.2% HV ± 0.3-0.6% HV	0,1 HR ± 0.3 HR	0.2 HR 0.6 HR
2a	Standard deviation, s	0.2% HB			
3	Error limits at 0.95 confidence level, δ Standard deviation of calibrated value, s	0.4-0.9% HB	0.15-0.25% HV	0.15-0.20 HR	0.3-0.5 HR
3a	Standard deviation, s	0.5% HB			
4	Error limits at 0.95 confidence level, δ Standard deviation of calibrated value, s	1.6-2.1% HB	0.3-0.8% HV	0.3-0.5 HR	0.5-1.2 HR
4a	Standard deviation, s	1% HB			
5	Maximum permitted error, Δ	4-5% HB	3-5% HV	± 1.2 HRA ± 2.0 HRB ± 1.0-2.0 HRC	± 1-2 HRN ± 2-3 HRT

In machinery production where hardness testing is widely employed, the smallest units of mass and length which are necessary in practice are in the order of 10^{-6} kg and 10^{-6} m, respectively. The uncertainty of the international and national standards of these units is 10^{-9} or better. This "reserve" of three orders of magnitude is sufficient for the realization of a traceability in several steps. When stating the hardness of a workpiece, values used to be indicated in steps of 0.5 HRC. By employing the classical methods of metrology, hardness testing machines having an uncertainty of 0.1 HRC would be necessary to this end, what is practically impossible, since only standardizing machines have this uncertainty. If a denotes the uncertainty of ordinary measuring instruments and the standard is n-times better, the composed uncertainty is found to be

$$a' = \sqrt{a^2 + a^2n^{-2}} = a\sqrt{1 + n^{-2}}$$

The values for different ratios n are shown in Table 5.

Table 5

	n = 10	5	2
$\sqrt{1 + n^{-2}}$	1.005	1.020	1.118
(a' - a) %	0.5	2.0	11.8

A ratio n = 10 between uncertainties at two levels of the hierarchy scheme is the well known thumb rule of metrology, but it can be reduced to 5 if necessary. In the hierarchy schemes of hardness measurement n = 2 is usual. Consequently the effect of the higher levels may augment uncertainties by 12 % what is not negligible any more.

A standard or recommendation on questions of the accumulation of uncertainties in the field of hardness measurement would be useful and desirable.

In the followings a calculation method [P-24] is given for the propagation of random uncertainties in the hierarchical order shown in Fig. 5. (This scheme is similar to the one shown in Fig. 2). The hardness value is transferred from standardizing machine 1 to hardness testing machine 3 by standardized blocks 2. The test specimen (workpiece) is designated by 4. The three measurement operations are denoted by A, B, and C. In the case of most measured quantities the traceability of an ordinary measuring instrument to the standard generally necessitates only two measurements. In the case of hardness testing three measurements are required. Each additional level in the hierarchy scheme results in an increase of uncertainty.

Uncertainties are classified, according to CIMP Recommendation 1 (CI-1981) in two groups. Those which are determined by statistical methods are characterized by the standard deviation s. The second group of uncertainties includes the ones which are determined by other, not statistical methods. These methods inevitably include subjective factors (estimate, etc.). Uncertainties belonging to this group are characterized

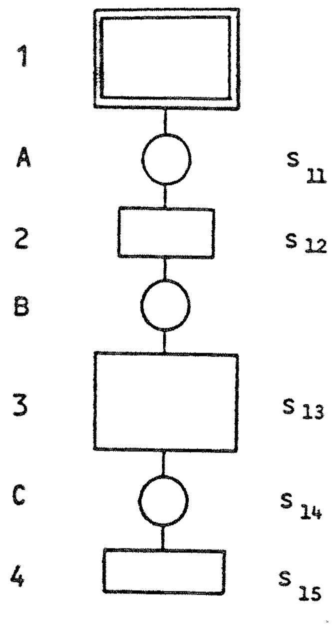

Fig. 5.
Hierarchy scheme

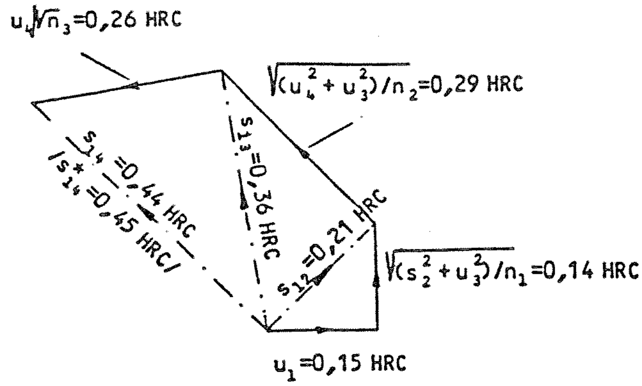

Fig. 6.

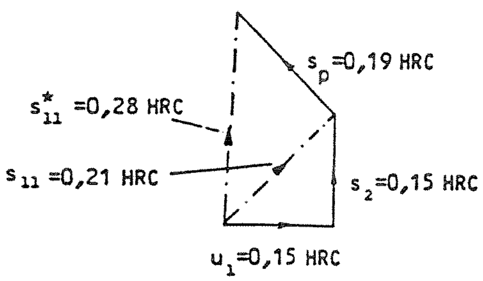

Fig. 7.

a)

A	B	C	D

b)

A	B	C	D
B	D	A	C
C	A	D	B

c)

A	B	C	D
B	D	A	C
C	A	D	B
D	C	B	A

Fig. 8. Latin Square developed from randomized blocks.

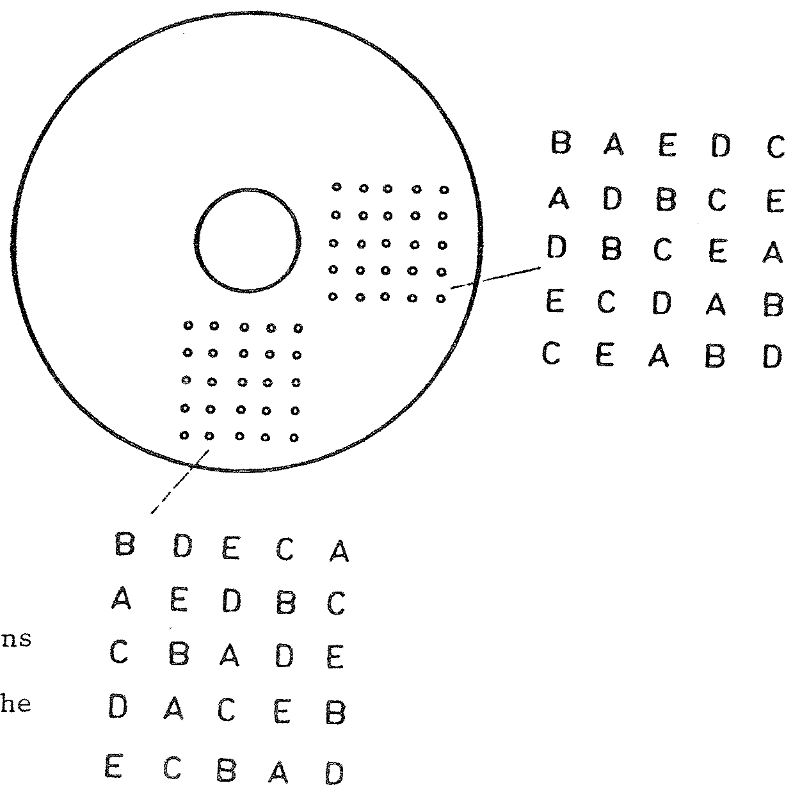

Fig. 9. Hardness testing indentations arranged in two 5 x 5 Latin Squares on the surface of the test block.

by a value $\underline{u}$, which is treated as if it were a standard deviation. At calculating and specifying random uncertainties of hardness testing the following symbols will be used.

Ramdom uncertainties

u_1 variation of the standardizing machine in time
s_2 repeatibility of the standardizing machine
u_3 non-uniformity of hardness along the surface of blocks
u_4 repeatability of the hardness testing machine
u_5 non-uniformity of hardness of the specimen
u_6 variation of the hardness testing machine in time.

Number of indentations

n_1 at calibrating standardized blocks
n_2 at calibrating hardness testing machines
n_3 at hardness testing
n_o at each machine, when comparing two standardizing machines.

To the determination of the individual values [P-23, I-5, P-5, P-28]:
To s_2. For the most important hardness testing methods,

$$s_2 = 0.003\ e \quad \text{HRC},$$

where $\underline{e}$ denotes the permanent increase of depth of indentation in HRC units, further

$$s_2 = (4.5\ \text{HV} + 300)\ .\ 10^{-3} \quad \text{HV 30},$$

and

$$s_2 = (1 + \text{HB}/600) \qquad \text{HB 10/3000}.$$

If there were standardized blocks with perfectly uniform surface, the standardizing machine would measure it with repeatability standard deviation s_2. On real blocks local hardness non-uniformity can be separated by analysis of variance. In Table 6, s_2, u_3 and u_4 are given for two hardness levels as examples.

To u_1: This is a value estimated on the basis of long-time experiments.

According to measurements on a HRC standardizing machine

$$u_1 \cong s_2.$$

To u_3: An upper limit for this estimated value can be obtained from standards, from the value specified for the range of five indentations. (Standard deviation s_R estimated from range R is found to be

$$s_R = A(n)\ R$$

where

$$A\ (5) = 0.43 \quad \text{and}$$
$$A(10) = 0.32 \qquad)$$

In the case of Rockwell C blocks the specified value for the range of five indentations is 1.5 % of $\underline{e}$. From this an estimate for the corresponding standard deviation can be calculated. After having

Table 6

Hardness level	s_2	u_3	u_4
30 HRC 65 HRC	0.2 0.1	0.4 (0.22) 0.2 (0.11)	0.8 0.4
200 HV 800 HV	1.2 4.0	3.3 9.5	6.2 18.3
100 HB 300 HB	1.2 1.5	1.6 2.6	3.7 5.4

Table 7

	S R V R V R V S
ST	$s_{11}^2 = u_1^2 + s_2^2$
B	$s_{12}^2 = u_1^2 + (s_2^2 + u_3^2)/n_1$
HT	$s_{13}^2 = u_1^2 + (s_2^2 + u_3^2)/n_1 + (u_4^2 + u_3^2)n_2$
M_i	$s_{14}^2 = u_1^2 + (s_2^2 + u_3^2)/n_1 + (u_4^2 + u_3^2)/n_2 + u_4^2/n_3$
M_r	$s_{15}^2 = u_1^2 + \underbrace{(s_2^2 + u_3^2)}_{A}/n_1 + \underbrace{(u_4^2 + u_3^2)}_{B}/n_2 + (u_4^2 + u_5^2)/n_3 + u_6^2$
ST_1-ST_2	$s_{o1}^2 = 2\,[u_1^2 + (s_2^2 + u_3^2)/n_o]$

separated s_2, the value of u_3 is obtained. In Table 6 the values for selected HRC blocks with reduced range of 1 % are also included in brackets.

The correlations of u_3 and s_2 deserve a closer consideration. Metrological development work aimed at improving the uniformity of hardness measurement has been carried in two directions, namely at improving the metrological characteristics of the standardizing machines and of the blocks, respectively. Efforts are required in both directions in parallel. When measuring a block on the standardizing machine, random uncertainties s_2 and u_3 cannot be separated, only their sum appears. Even if one of the two uncertainties were equal to zero, one could not detect this fact. But if the two values are considerably different, their sum is practically equal to the greater one. What is the situation with the actually available equipment?

As Table 6 shows, the standard deviation u_3 for the non-uniformity of a block having a range of 1.5 % is approximately the double of the standard deviation of the standardizing machine (s_2). By the analysis of variance, non-uniformity of hardness can be shown to be significant. In the case of blocks with a range of 1 %, however (see values in brackets in Table 6), values s_2 and u_3 are approximately equal. Manufacturers of standardizing machines and of blocks arrived to a similar level of random uncertainties. A simple calculation shows that s_2 = 0.003 e corresponds, in the case of 5 indentations, to a range of 0.7 % of e. Consequently if somebody wanted to narrow the specified range at the calibration of blocks to the value of 0.7 %, this specification could be met only on a perfectly uniform block. Or the standardizing machine should be improved. Indeed, at measurements performed on a standardizing machine with laser interferometer ranges of 0.5-0.6 % were found on the blocks of three different manufacturers, and ranges of 0.7-0.8 % for those of other three sources [C-9].

The development of standardizing machines and blocks can mutually promote or retard the development of the other, or already achieved levels of development can be detected by measurement.

To u_4: An upper limit for this estimate is given by the range of five indentations specified in standards as repeatability. From this an estimate of the standard deviation is calculated, from which u_3 is still to be separated. See the examples in Table 6. The hardness testing machine would measure a perfectly uniform test specimen with this standard deviation, or better.

To u_5: This value is included here only for the sake of completeness. It actually belongs to the sphere of testing of materials.

To u_6: It is difficult to determine a value for the variance of the hardness testing machine in time. During calibration a correction with respect to the national hardness reference scale can be established. This correction is changing with time. The user of the apparatus can determine variations of the testing machine in time by daily calibrations by means of standardized test blocks and plot them in a chart. This value should be determined, if required, individually for each machine.

Accumulation of random uncertainties

Calculation is carried out by the quadratical addition of standard deviations, that is by the addition of variances. Repeated quadratical addition can be demonstrated by right-angled triangles (Fig. 6, 7). The addition of variances accumulated in hardness measurement is shown in Table 7, in the modified clear presentation proposed by Iizuka [I-5].

Accumulated standard deviations figuring in Table 7:

s_{11} Hardness standardizing machine, or the national hardness reference scale, a single measurement.

s_{12} Nominal value of the calibrated reference block (B).

s_{13} Calibration of the hardness testing machine (HT).

s_{14} Hardness measurement by the hardness testing machine shortly after calibration (M_1).

s_{15} Hardness measurement by the hardness testing machine, taking into consideration also its variation in time and the non-uniformity of the specimen (M_r).

s_{01} Comparison of two standardizing machines, i.e. the standard deviation of the difference of two hardness reference scales (ST_1-ST_2).

In Fig.5 the places are marked where these values are obtained. This presentation of error accumulation clearly shows that a term for repeatability (R) and for non-uniformity of the specimen (V) is added at each level of the hierarchy scheme. At the first and last level a term representing stability in time (S) is also added. A denotes the sum corresponding to the range specified in standards for non-uniformity of blocks. Similarly B corresponds to repeatability of the testing machine as specified in standards.

As numerical examples, the values calculated for 50 HRC according to Table 7 are shown in Fig 6. Random uncertainty s_{11} of the standardizing machine is shown in Fig. 7. If the unknown systematic uncertainties s_p of the standardizing indenter are also taken into consideration, s_{11} = 0.21 HRC increases to $s_{11}*$ = 0.28 HRC. In the addition according to Fig. 6 s_p has only a minor effect: s_{14}^* is 0.01 HRC higher than s_{14}. By considering that several values used in this calculation are estimates, this difference is negligible.

A broken line divides Table 7. What is above this line, is the task of the Institutes of metrology: Standardized blocks ensuring the traceability of hardness testing machines to the national hardness reference scale are placed at the disposal of material testing laboratories. Random uncertainties figuring below the broken line are less reliable, informative values.

International hardness comparisons

It is to be hoped that generally accepted international standard machines will be available for the hardness scales in the not very distant future (Fig. 1).

The correlation between different national standardizing machines to-day is ensured by international hardness comparisons (Fig. 2). Random uncertainty of a comparison is given in Table 7 (variance s_{01}^2). The hardness comparison is actually a standardization of the same block of high quality (non-uniformity of 1 %) by an increased number of indentations on two standardizing machines. A calculation gives e.g. s_{01} = 0.30 HRC at 30 HRC, and 0.15 HRC at 65 HRC, if n_o = 25. The formula for s_{01} is analogous to that for s_{12}. A closer analysis of the formula shows, that practically

$$s_{01}^2 = 2\,u_1^2$$

The uncertainty of the comparison, the reliability of systematic differences or corrections determined by comparison depend solely on long-time stability of the two standardizing machines. Other influence factors (repeatibility, non-uniformity of the block) can be nearly neglected in consequence of the increased number of indentations employed for comparison purposes.

Systematic errors, corrections

The calculation of the propagation and accumulation of uncertainties discussed so far, as summarized in Table 7, deals only with random uncertainties. At various levels of the hierarchy scheme systematic errors may also arise, these can however be eliminated by employing a correction to the hardness value, equal to the systematic error with opposite sign.

When tracing back a hardness value to the measurement standards, systematic errors may arise at the following levels.

- The hardness testing machine may have an error with respect to the block used for its calibration (e.g. permitted maximum error of ± 1.5 HRC according to ISO 716).

  Unfortunately, this error is not sufficiently stable to be taken into consideration with opposite sign as a correction.

- The blocks are supposed to be marked in values of the national hardness reference scale, when calibrated on the national standardizing machine.

- The national standardizing machine may have a systematic difference with respect to similar standards of other states. This fact should not be forgotten, if there is a danger of an international dispute on the hardness of a component. (This problem is discussed more in detail in the Chapter 5 dealing with international comparisons).

3. Experimental methods of uncertainty evaluation

3.1. The problem of the true value of the measured object

In the metrological examination of a measuring instrument we need the true value of the quantity to be measured (VIM 1.18). The true value of a quantity is an ideal concept and, in general, we cannot know it exactly. For a given purpose, the conventional true value of the quantity (VIM 1.19) may be substituted for the true value.

In the case of hardness even the determination of the conventional true value of a measured object may be problematic.

One source of the problems is the non-repeatability of hardness measurement. In the case of the measurement of other quantities we can come nearer to the conventional true value by repeating the measurement several times. The hardness measurement causes a local destruction of the surface of the measured object, it cannot be repeated at the same place. Adjacent points of the measured object have not necessarily the same hardness even in their original state. But the deformation caused by the previous indentations in that region of the surface further deforms the original state of hardness.

Consequently, when we speak of repeated hardness measurement (e.g. n = 5 measurements of the same object) we have not measured the same quantity.

In the BIML Publication Hardness test blocks and indenters [P-22] (point 4.1), measurement results were presented showing some characteristic hardness value distributions along the surface of a test block. If we want to determine the repeatability of a hardness testing equipment, the obtained repeatability value inevitably includes also the non-uniformity of hardness of the block, beside the repeatability of the measuring instrument. This is not the case e.g. when measuring the mass, since the measured quantity, the measured mass can be assumed to be constant, to have its conventional true value.

In analysing the uncertainty of hardness measurement some methods are to be found to separate the uncertainties of the measuring instrument from those of the measured object.

The factorial experiment employed by YAMAMOTO, YANO and YAZIMA [Y-4, Y-5, Y-6] was already summarized in [P-22].

The method of Latin Squares, first employed on hardness blocks by MARRINER [M-1] deserves a detailed discussion [P-11, P-7, P-18, P-22].

3.2. The experimental method of Latin and Greco-Latin Squares

Similarly as various problems of probability calculation and combinatorics, the latin and greco-latin squares were first discussed in connection with gambles and puzzles. The so-called magic square of n elements is a square composed of the numbers 1,2,... n in such a way that each number figures only once in each row and column. E.g. a magic square of four elements is

$$\begin{matrix} 3 & 4 & 1 & 2 \\ 1 & 3 & 2 & 4 \\ 4 & 2 & 3 & 1 \\ 2 & 1 & 4 & 3 \end{matrix} \qquad (1)$$

The sum of each row and column is equally n (1 + n)/2 = 10. A problem described in this connection by Euler in 1779 was named the problem of the 36 officers. As an introduction to the Latin Squares it is worth while to mention this problem, or rather than of the 25 officers:

At a military parade there are 25 officers from 5 different brigades (α, ß, γ, δ, ε). From each brigade one lieutenant (A), one captain (B), one major (C), one colonel (D), and one brigadier (E).

The first task is to arrange the 25 officers in a 5 x 5 square so that each grade be represented once in each row and column. This is possible by arranging the officers in a five-element magic square, a Latin Square, e.g.

$$\begin{matrix} B & C & D & E & A \\ C & D & E & A & B \\ D & E & A & B & C \\ E & A & B & C & D \\ A & B & C & D & E \end{matrix} \qquad (2)$$

A further contraint can be that not only each grade, but also each brigade should be represented in each row and column, but only once. In this case the arrangement in a Greco-Latin Square can help, e.g.

$$\begin{matrix} B\beta & C\gamma & D\delta & E\varepsilon & A\alpha \\ C\varepsilon & D\alpha & E\beta & A\gamma & B\delta \\ D\gamma & E\delta & A\varepsilon & B\alpha & C\beta \\ E\alpha & A\beta & B\gamma & C\delta & D\varepsilon \\ A\delta & B\varepsilon & C\alpha & D\beta & E\gamma \end{matrix} \qquad (3)$$

The Greco-Latin Square is produced by superimposing two Latin Squares and using Greek letters in one of them for the sake of discrimination. Two Latin Squares are called orthogonals if in the resulting Greco-Latin Square each couple of Latin and Greek letters figures only once. On the contrary the following combination of two squares is non orthogonal.

$$\begin{matrix} A & B & C \\ C & A & B \\ B & C & A \end{matrix} + \begin{matrix} \alpha & \gamma & \beta \\ \beta & \alpha & \gamma \\ \gamma & \beta & \alpha \end{matrix} = \begin{matrix} A\alpha & B\gamma & C\beta \\ C\beta & A\alpha & B\gamma \\ B\gamma & C\beta & A\alpha \end{matrix} \qquad (4)$$

In the Greco-Latin Square resulting from the superimposition of the two Latin squares the couples Aα, Bγ and Cß occur three times each, while 6 other possible couples of letters are missing. Such a square is useless for experimental purposes. Square (3) proves that the problem of 25 officers can be solved. Euler proved that there are no 6 x 6, or 10 x 10 element orthogonal squares, the problems of 36 or of 100 officers cannot be solved.

From a given square several further squares can be generated by permutating the rows and columns. By a similar operation squares can be reduced to the form where the first row and first column contain the ordered series of letters. E.g. The first square in (4) can be reduced by interchanging the 2nd and 3rd rows to the form

```
A B C
B C A
C A B          (5)
```

The number of possible reduced five element squares is 56, that of six element squares is 9408. From these, further squares can be generated by permutating rows or columns. From a reduced Latin Square of n elements $(n!)^3$ further squares can be generated.

To ensure random arrangement for an experiment, random numbers taken from a table can be used for permutation of rows and columns. E.g. the randomization of a four element reduced Latin Square by the help of random numbers 1342 and 2134 is the following

	1.	2.	3.	4.
1.	A	B	C	D
2.	B	D	A	C
3.	C	A	D	B
4.	D	C	B	A

(6)

	1.	3.	4.	2.
1.	A	C	D	B
2.	B	A	C	D
3.	C	D	B	A
4.	D	B	A	C

(7)

	1.	3.	4.	2.
2.	B	A	C	D
1.	A	C	D	B
3.	C	D	B	A
4.	D	B	A	C

(8)

Three to nine element Latin Squares can be found in Fisher, R.A.-Yates, F.: Statistical Tables for Biological, Agricultural and Medical Research, London, 1957.

Like many other methods of factorial experiment also the Latin Squares were first employed in biological and agricultural research.

An example which gave the idea to MARRINER to use the method for hardness test blocks is illustrated in Fig. 8. The yields of four types of wheat are to be compared. The simplest solution is to sow the four types of wheat (A,B,C,D) in adjoining plots (Fig. 8a). Yields may however be influenced by different soil quality in the different plots. Reliability of the experiment can be improved by repeating the experiment with a different arrangement in other blocks (Fig. 8 b). A further improvement can be brought about by arranging the plots in a Latin Square (Fig. 8 c). E.g. if soil quality is improving as we pass from the left to the right, this change of soil quality has a similar effect on the yield of each wheat type, since each figures once in each column.

The analogy is valid also for the hardness along the surface of the hardness test block. For the determination of the systematic errors (corrections) of several indenters, 5 to 10 indentations are made by each of them, as well as by the standard indenter on the same block, on a standardizing machine. The mean of hardness values obtained by one indenter is used for calculating the correction.

In the case of good quality indenters correction values obtained by this method may be lower than the variation of values obtained by the same indenter. Hardness value may be influenced by the location of the indentation on the block. But hardness distribution on the surface of the block is not known in advance. Latin Squares can help in such cases, just as in the wheat crop experiment described above.

If we want to measure in five rows, five columns with five indenters, the number of necessary combinations would be $5^3 = 125$. But there are only $5^2 = 25$ points available, on account of the non-repeatability of the hardness test at the same point. So the use of a Latin Square not only reduces the number of measurements from 125 to 25, but this is the only possible solution to detect the effect of the surface in two co-ordinate directions and of different indenters.

The actual realization of the Latin Square on the surface of the block can be a grid with a division of 3.5 mm. So a 5 x 5 square takes a block surface of 14 x 14 mm. Hardness variations on such a small surface can be proved by a Latin Square experiment evaluated by the analysis of variance.

Accordingly, in the experimental design of Latin Squares the effect of three factors is examined.

- Different indenters.
- Hardness variation in horizontal direction.
- Hardness variation in vertical direction.

(As the block surface has two dimensions, hardness variations of the block are considered as two factors.)

A single measured hardness value, can be represented by the mathematical model

$$\bar{x}_{ij} = \bar{x} + p + h + v \pm s \qquad (9)$$

where

- x_{ij} is the hardness value obtained in the ith row and jth column,
- $\bar{x}$ is the mean hardness of the block,
- p is the systematic error of the indenter used,
- h is the systematic error brought about by hardness variation in the horizontal direction,
- v is the same in the vertical direction,
- s is measurement uncertainty.

Measurement results obtained in a Latin Square (with the symbol of the respective indenter in brackets) can be presented as shown in (10). This arrangement corresponds to a map of the examined surface of the block.

$$x_{11}(B)\ x_{12}(D)\ x_{13}(E)\ x_{14}(C)\ x_{15}(A)$$

$$x_{21}(A)\ x_{22}(E)\ x_{23}(D)\ x_{24}(B)\ x_{25}(C)$$

$$x_{31}(C)\ x_{32}(B)\ x_{33}(A)\ x_{34}(D)\ x_{35}(E)$$

$$x_{41}(D)\ x_{42}(A)\ x_{43}(C)\ x_{44}(E)\ x_{45}(B)$$

$$x_{51}(E)\ x_{52}(C)\ x_{53}(B)\ x_{54}(A)\ x_{55}(D) \qquad (10)$$

Fig. 9 shows a block with two different Latin Squares of 5 x 5 indentation. One of the squares is the same as the one given in (10).

The measurement results x_{ij}, if developed according to (9) without the uncertainty component, can be written for the Latin Square as shown in (11). Sums of rows X_{vi} and sums of columns X_{hj} are given to the right and below the square respectively. Sums of values measured by the same indenter X_{pk}, collected from the square, are also given below.

A comparison of sums of rows X_{vi} shows that each one contains the component

$$5\,\bar{x} + \sum_{A}^{E} p + \sum_{1}^{5} h_j,$$

i.e. the mean hardness of the block, the systematic errors of the indenters and hardness variation in the horizontal direction.

Sums of rows differ only in the components $5v_1\ 5v_2 \ldots 5v_5$, representing the systematic variation of each row caused by the variation of hardness in the vertical direction. Accordingly the variation of the sums of row X_{vi} is characteristic of the variation of hardness in one direction.

Analogous considerations show that the variation of hardness in the horizontal direction is characterized by the variation of sums of colums X_{hj}. Finally the variation of sums obtained with the same indenter X_{pk} characterizes the variation brought about by the use of different indenters, independently of hardness variations on the block surface.

These consideration apply only in the case if there are no interactions between factors. In the relatively small surface of the square, independence of hardness variations in the two co-ordinate directions can be assured, supported by measurement results. The correction value of indenters, in turn, is a function of hardness value. In the case of a Latin Square, however, hardness variation is so small that it does not influence the correction of indenters.

$$
\begin{array}{llcl|l}
(\bar{x}+p_B+h_1+v_1) & (\bar{x}+p_D+h_2=v_1) & \cdots\cdots\cdots & (\bar{x}+p_A+h_5+v_1) & X_{v1} = 5\,\bar{x} + \sum_A^E p + \sum_1^5 h_j \boxed{+\,5\,v_1} \\
(\bar{x}+p_A+h_1+v_2) & (\bar{x}+p_E+h_2+v_2) & \cdots\cdots\cdots & (\bar{x}+p_C+h_5+v_2) & X_{v2} = 5\,\bar{x} + \sum_A^E p + \sum_1^5 h_j \boxed{+\,5\,v_2} \\
\vdots & \vdots & & \vdots & \vdots \\
(\bar{x}+p_E+h_1+v_5) & (\bar{x}+p_C+h_2+v_5) & \cdots\cdots\cdots & (\bar{x}+p_D+h_5+v_5) & X_{v5} = 5\,\bar{x} + \sum_A^E p + \sum_1^5 h_j \boxed{+\,5\,v_5} \\
\hline
\end{array}
$$

$$
X_{h1} = 5\,\bar{x} + \sum_A^E p + \underline{5_{h1}} + \sum_1^5 v_i \qquad X_{h2} = 5\,\bar{x} + \sum_A^E p + \underline{5_{h2}} + \sum_1^5 v_i \; ; \ldots\ldots \; X_{h5} = 5\,\bar{x} + \sum_A^E p + \underline{5_{h5}} + \sum_1^5 v_i
$$

$$
\begin{aligned}
X_{pA} &= 5\,\bar{x} + \boxed{5p_A} + \sum_1^5 h_j + \sum_1^5 v_i \\
X_{pB} &= 5\,\bar{x} + \boxed{5p_B} + \sum_1^5 h_j + \sum_1^5 v_i \\
&\;\;\vdots \\
X_{pE} &= 5\,\bar{x} + \boxed{5p_E} + \sum_1^5 h_j + \sum_1^5 v_i
\end{aligned}
\qquad (11)
$$

In a Greco-Latin Square arrangement four factors can be examined. In hardness standardizing work the fourth factor may be e.g. the order of measurements performed by the same indenter. One may suppose that the indenter clamped into the standardizing machine changes its position at the effect of each indentation process even if dummy indentations were performed outside the square following a change of indenter.

A 7 x 7 element Greco-Latin Square is given in (12).

G2	B3	C6	E5	F1	A7	D4
F4	D7	G5	A3	C2	E1	B6
C7	E6	F2	B4	G3	D5	A1
A6	G1	D3	F7	E4	B2	C5
B5	A2	E7	C1	D6	G4	F3
D1	F5	A4	G6	B7	C3	E2
E3	C4	B1	D2	A5	F6	G7

(12)

For the sake of convenience, Greek letters were replaced here by numbers, which represent the order of the measurement made by the same indenter. While in the Latin Square shown in (10) the five indentations to be made by e.g. indenter D could be made in any order, in the Greco-Latin Square the indenter D is first used at D1 (6th line, first column), then at D2 (7th line, 4th column), and so on. The distance between indentations is 3.5 mm, the whole square embraces a surface of 21 x 21 mm.

The mathematical model for the Greco-Latin Square experiment is

$$x_{ij} = \bar{x} + p + h + v + o \pm s \qquad (13)$$

where

o is the variation due to the order of measurement,
all the other symbols correspond to (9).

From the square a fourth sum, X_{ol}, the sum of identical orders can be calculated, the variance of which is characteristic of the effect of the number of measurements performed after having changed the indenter (irrespective of the dummy indentations).

3.3. Analysis of variance

The great problem in the metrology of hardness is that the variance of a measured hardness value (s^2) is the sum of the variance of the non-uniformity of the measured object (in our case the hardness test block, s_B^2) and of the variance of the hardness tester (s_T^2).

$$s^2 = s_B^2 + s_T^2$$

It is the value s which can be obtained from measured hardness values. Before the introduction of the statistical method of variance analysis, it was mathematically impossible to decompose s to its components [P-14].

Measurement results obtained in the Latin Square arrangement are processed by this method of mathematical statistics, by the analysis of variance. The method is amply discussed in the relevant literature. Here only the evaluation process is described, without going into details, into mathematical justification.

To perform the analysis of variance of measurement data according to (11) we need the sums X_{pk}, X_{vi}, X_{hj}, as well as the squares thereof, namely X^2_{pk}, X^2_{vi}, X^2_{hj}.

Further the overall sum of the 25 values

$$X = \sum_{1}^{25} x_{ij} = \sum_{A}^{E} X_{pk} = \sum_{1}^{5} X_{vi} = \sum_{1}^{5} X_{hj},$$

and the square of this sum, X^2.

The tabulated form of the analysis of variance is given in (14), with calculation formulae.

The residual sum of squares X_r^2, can be determined by substracting the first three SS values from the overall value. s_1^2 is an estimation for the variance σ^2 characterizing the measuring equipment and measurement method. According to the column headed "Expectation of mean square", an estimation can be obtained for the variance between indenters, σ_p^2, for the variance between rows, σ_v^2 and between columns, σ_h^2.

These components can, however, be separated only, if the F-test proves that a component is significantly different from σ^2. The test values are indicated in the last column. The critical values of the F-test (F_{cr}) can be taken from statistical tables, for 95 %, or 99 % confidence levels. If the F values calculated according to the last column are higher than the $F_{cr,95}$ or $F_{cr,99}$ values,

$$F > F_{cr,95} \qquad (F > F_{cr,99}),$$

then it can be understood, with a confidence of 95 % (99 %), that the respective component of variance has a significant effect. In the case

$$F < F_{cr}$$

one cannot simply state the contrary. One can understand that on the basis of measurement results one cannot reject the hypothesis that the two variance values are equal.

After the F-test, the non significant sources are added to the residual and new F-tests are made on the modified table. In this case, however, new F_{cr} values apply on account of the modified degrees of freedom.

The main idea of variance analysis is summarized by the following formula.

$$f.s^2 = f_p.s_4^2 + f_v.s_3^2 + f_h.s_2^2 + f_r.s_1^2 \qquad (15)$$

Source	Sum of squares, SS	Degrees of freedom, f	Mean square, MS = SS/f	Expectation of mean square	Fisher test value F
Between indenters	$\frac{\sum_{A}^{E} X^2_{pk}}{5} - \frac{X^2}{25}$	$f_p = n - 1 = 4$	s_4^2	$s_4^2 \approx \sigma^2 + 5\,\sigma_p^2$	s_4^2/s_1^2
Between rows	$\frac{\sum_{1}^{5} X^2_{vi}}{5} - \frac{X^2}{25}$	$f_v = 4$	s_3^2	$s_3^2 = \sigma^2 + 5\,\sigma_v^2$	s_3^2/s_1^2
Between columns	$\frac{\sum_{1}^{5} X^2_{hj}}{5} - \frac{X^2}{25}$	$f_h = 4$	s_2^2	$s_2^2 = \sigma^2 + 5\,\sigma_h^2$	s_2^2/s_1^2
Residual	X_r^2	$f_r = 12$	s_1^2	$s_1^2 \approx \sigma^2$	1
Overall	$\sum_{1}^{25} x_{ijk} - \frac{X^2}{25}$	$f = 24$	s^2		

(14)

The overall sum of squares can be decomposed to the sums of squares originating from the different sources.

In the Latin Square experiment discussed here, the effect of three factors was examined. The Greco-Latin Square arrangement permits also the examination of a fourth factor, namely the order of measurements made by the same indenter. As already mentioned, it was supposed that the position of the clamped indenter is changing at the effect of each indentation, even if 2-3 unevaluated dummy indentations have been made after clamping the indenter.

For the analysis of variance we need the same sums, and the squares thereof, as in the case of Latin Square experiments, plus the sums of orders $X_{o\ell}$ (e.g. the sum of the third measurements made by each indenter), and $X^2_{o\ell}$.

The table of the analysis of variance is given under (16).

The calculation is similar as in the case of Latin Squares. The calculation method will be clarified by two numerical examples.

3.4. Numerical examples

Example 1. Latin Square of 5 x 5 elements

In a Latin Square experiment according to (10) with five Rockwell C indenters on a block of about 50 HRC the following results were obtained:

						X_v	HRC
	36(B)	31(D)	18(E)	-6(C)	18(A)	97	49.51
	-1(A)	13(E)	28(D)	27(B)	7(C)	74	49.63
	-3(C)	16(B)	0(A)	20(D)	-5(E)	28	49.86
	17(D)	-7(A)	-8(C)	-1(E)	11(B)	12	49.94
	20(E)	-7(C)	16(B)	0(A)	6(D)	35	49.83
X_h	69	46	54	40	37	X = 246	
HRC	49.65	49.77	49.73	49.80	49.32	HRC 49.75	

Sum of results obtained with the same indenter:

	A	B	C	D	E
X_p	10	106	-17	102	45
HRC	49.95	49.47	50.09	49.49	49.78

The HRC values given in the table serve only as information, but are not necessary for the analysis of variance.

The table contains the transformed measured values

$$e' = (e - 50) \cdot 40$$

where e is the permanent increase of depth of indentation in units of 2 µm (see ISO 6508-1986). e' = 40 corresponds to 1 HRC.

Source	Sum of squares, SS	Degrees of freedom, f	Mean square, MS = SS/f	Expectation of mean square	Fisher test value F
Between orders	$\frac{\sum_{\ell=1}^{7} X_{o\ell}^2}{7} - \frac{X^2}{49}$	$f_o = n - 1 = 6$	s_5^2	$s_5^2 \approx \sigma^2 + 7\,\sigma_o^2$	s_5^2/s_1^2
Between indenters	$\frac{\sum_{k=A}^{G} X_{pk}^2}{7} - \frac{X^2}{49}$	$f_p = 6$	s_4^2	$s_4^2 = \sigma^2 + 7\,\sigma_p^2$	s_4^2/s_1^2
Between rows	$\frac{\sum_{i=1}^{7} X_{vi}^2}{7} - \frac{X^2}{49}$	$f_v = 6$	s_3^2	$s_3^2 = \sigma^2 + 7\,\sigma_v^2$	s_3^2/s_1^2
Between columns	$\frac{\sum_{j=1}^{7} X_{hj}^2}{7} - \frac{X^2}{49}$	$f_h = 6$	s_2^2	$s_2^2 \approx \sigma^2 + 7\,\sigma_h^2$	s_2^2/s_1^2
Residual	X_r^2	$f_r = 24$	s_1^2	$s_1^2 \approx \sigma^2$	1
Overall	$\Sigma x_{ij}^2 - \frac{X^2}{49}$	$f = 48$	s^2		

(16)

The transformation of measured values simplifies calculations done without a computer program. The reasoning underlying the transformation is the following.

On the measuring microscope indentation depth values were read to 0.1 µm. Since two observers were working in parallel, the mean of their values could have the lowest digit of 0.05 µm. E.g. the indentation made in the first row, second column, by indenter D was measured by the two observers as 101.6 and 101.5 µm, respectively. The mean was 101.55 µm. From all mean values 100 was substracted and the remainder (1.55) multiplied by 20, to obtain in this case 31. By this transformation both values less than one and higher than 100 were avoided.

Values necessary for the analysis of variance can be taken from the table:

$$\frac{\sum_{k=A}^{E} X_{pk}^2}{5} = 4810.80$$

$$\frac{\sum_{i=1}^{5} X_{vi}^2}{5} = 3407.60$$

$$\frac{\sum_{j=1}^{5} X_{hj}^2}{5} = 2552.40$$

$$\frac{X^2}{25} = 2420.64$$

$$\Sigma\, x_{ij}^2 = 6628.00$$

The table of the analysis of variance can be composed as follows.

		f	s^2	F
Indenters.....	4810.80 - 2420.64 = 2390.16	4	597.54	$10.3>F_{cr}$ (99)
Rows	3407.60 - 2420.64 = 986.96	4	246.74	$4.2>F_{cr}$ (95)
Columns	2552.40 - 2420.64 = 131.76	4	32.94	$0.6<F_{cr}$ (99)
Residual	4207.36 - 3508.88 = 698.48	12	58.21	1
Overall	6628.00 - 2420.64 = 4207.36	24	175.31	

From statistical tables we obtain for 4/12 degrees of freedom that

$F_{cr,95}$ = 3.26, at 95 % confidence level and
$F_{cr,99}$ = 5.41, at 99 % confidence level.

The F value for the effect between indenters is found to be significant at 99 % level, that for the effect between the rows at 95 % level. The effect between columns is not significant, consequently this source can be unified with the residual. The modified table is

		f	s^2	F
Indenters........	2390.16	4	597.54	11.5>F_{cr} (99)
Rows	986.96	4	246.74	4.8>F_{cr} (99)
Residual	830.24	16	51.89	1
Overall	4207.36	24	175.31	

The critical F values for 4/16 degrees of freedom:

$$F_{cr,95} = 3.01$$
$$F_{cr,99} = 4.77$$

In the modified table both sources, those of different indenters and rows are significant at the 99 % level.

The conclusions of this Latin Square experiment can be summarized as follows:

1. Estimation for the uncertainty of the measuring method and equipment (for designations see (14)):

$$\sigma^2 = 51.89$$

$$\sigma = 7.2 \mathrel{\widehat{=}} \pm 0.18 \text{ HRC}$$

2. Estimation of the standard deviation of mean values measured by different indenters:

$$\sigma_p^2 = \frac{s_4^2 - \sigma^2}{5} = \frac{597.54 - 51.89}{5} = 109.13$$

$$\sigma_p = 10.4 \mathrel{\widehat{=}} + 0.26 \text{ HRC}$$

It is not surprising that this effect is significant, since the establishment of differences between different indenters was the main objects of the experiment. σ_p characterizes the "homogeneity" of the group of indenters.

3. Estimation of the standard deviation of mean values obtained in different rows:

$$\sigma_v^2 = \frac{s_3^2 - \sigma^2}{5} = \frac{246.74 - 51.89}{5} = 38.97$$
$$\sigma_v = 6.2 \mathrel{\widehat{=}} \pm 0.15 \text{ HRC}$$

There is a noticeable, significant hardness variation on the block surface in the vertical direction.

4. There is no significant difference between the mean values obtained in different columns. The obtained F-value is so far from the critical value, that a hardness variation on the block surface in the horizontal direction is improbable, even in the case if interactions were not taken into consideration.

Example 2. Greco-Latin Square of 7 x 7 elements

Measurement results obtained on a block of approximately 65 HRC in the experimental arrangement according to (12) were as follows:

								X_v
	32	22	18	32	32	39	47	222
	4	38	44	41	18	35	41	221
	13	24	22	27	40	28	47	201
	35	31	30	22	28	33	26	205
	21	30	25	31	39	43	34	223
	10	4	40	37	28	27	38	184
	10	10	29	41	59	37	61	247
X_h	125	159	208	231	244	242	294	X=1503

The sums of orders and of indenters, respectively:

	1.	2.	3.	4.	5.	6.	7.
X_o	215	214	204	199	214	231	226

	A	B	C	D	E	F	G
X_p	291	201	143	233	192	155	288
HRC	63.96	64.28	64.49	64.17	64.31	64.45	63.97

These tables contain the transformed measurement results:

$$e' = (e - 35) \cdot 40$$

Values necessary for the analysis of variance:

$$\frac{\sum_{\ell=1}^{7} X_{o\ell}^2}{7} = 46\,210.14$$

$$\frac{\sum_{k=A}^{G} X_{pk}^2}{7} = 49\,093.28$$

$$\frac{\sum_{i=1}^{7} X_{vi}^2}{7} = 46\,449.28$$

$$\frac{\sum_{i=1}^{7} X_{hj}^2}{7} = 48\ 866.72$$

$$\frac{X^2}{49} = 46\ 102.22$$

$$\Sigma x_{ij}^2 = 53\ 375.00$$

With these values, the table of the analysis of variance was found to be:

		f	s^2	F
Orders	46 210.14 - 46 102.22= 107.88	6	17.98	$0.4<F_{cr}$
Indenters	49 093.28 - 46 102.22=2 991.06	6	498.51	$11.3>F_{cr}$
Rows	46 449.28 - 46 102.22= 347.06	6	57.84	$1.3<F_{cr}$
Columns	48 866.72 - 46 102.22=2 764.50	6	460.75	$10.4>F_{cr}$
Residual	1 062.28	24	44.26	1
Overall	53 375.00 - 46 102.22=7 272.78	48		

The critical F values for 6/24 degrees of freedom, as taken from statistical tables, are

$$F_{cr,95} = 2.51$$

$$F_{cr,99} = 3.67$$

The hypothesis of the experiment about the effect of the order of measurements was not found to be true, the F value for the variance factor of the order of the measurement is far below the critical value. The effect of rows is similarly not significant.

The effect of columns, however, is very strong, similar to that of indenters. By uniting the not significant factors with the residual, the modified table is found to be:

		f	s^2	F
Indenters........	2991.06	6	498.51	11.8
Columns..........	2764.50	6	460.75	10.9
Residual	1517.22	36	42.15	1
Overall	7272.78	48		

By comparing the new F-values, with the critical value for 6/36 degrees of freedom

$$F_{cr,99} = 3.35,$$

the previous statement on the two significant factors is confirmed.

The conclusions of this Greco-Latin Square experiment:

1. Estimation for the uncertainty of the measuring method and equipment:

$$\sigma^2 = 42.15$$

$$\sigma = 6.5 \mathrel{\widehat{=}} \pm\ 0.16 \text{ HRC}$$

This is conform with the ± 0.18 HRC value found in Example 1.

2. Estimation of the standard deviation of mean values measured by different indenters:

$$\sigma_p^2 = \frac{s_4^2 - \sigma^2}{7} = \frac{498.51 - 42.15}{7} = 65.2$$

$$\sigma_p = 8.1 \mathrel{\widehat{=}} 0.20 \text{ HRC}$$

This value depends on the composition of the group of indenters used in the experiment. The σ_p value is not especially interesting from metrological point of view, since the systematic errors of individual indenters determined in the experiment are used as correction values in later measurements.

3. Estimation of the standard deviation of mean values obtained in different columns:

$$\sigma_h^2 = \frac{s_2^2 - \sigma^2}{7} = \frac{460.75 - 42.15}{7} = 59.8$$

$$\sigma_h = 7.7 \mathrel{\widehat{=}} \pm\ 0.19 \text{ HRC}$$

There is a significant hardness variation on the block surface in the horizontal direction, which is similar to that observed in Example 1 in the vertical direction.

4. There is no significant difference between rows and between the orders of measurement.

3.5. Some remarks on experiments in Latin or Greco-Latin Squares

Experiments in the square arrangement are especially useful in the first stage of a longer series of experiments, when it is not yet known which factors have a significant effect. Later the non significant factors can be disregarded and a simpler arrangement is sufficient. E.g. Latin Square or factorial experiment in place of the Greco-Latin Square.

A disadvantage of the square arrangements is that all factors should occur at the same number of levels. E.g. in Example 2 seven indenters, seven measurements with each, in seven rows and seven columns. In experimental practice the use of squares with 5 to 9 elements is usual. In the case of a higher number of elements the calculation may become complicated (if no computer program is disponible) and it is difficult to find identical levels for each factor. In the case of 4 elements, in turn, random distribution is not sufficiently ensured.

Therefore in such a case it is advisable to perform two similar experiments, in two different arrangements. Another disadvantage of four element squares is that the number of freedoms remaining for the residual, namely the measuring method, is low.

To ensure random distribution, it is advisable to avoid regular squares, in which rows are generated by shifting the elements of the preceding row without varying the order of the letters, like the squares given in (4).

In exceptional cases one of the <u>values</u> of the square may be <u>missing</u>. E.g. an external influence disturbs the measurement, or anyhow an evidently outsider value would necessitate the repetition of the complete series of measurement. (Hardness measurements cannot be repeated at the same place!)

A single missing value can be replaced by an estimation taken from neighbouring values, by using the following formula in the case of Latin Squares:

$$x' = \frac{n\ (X_p' + X_v' + X_h') - 2\ X'}{(n - 1)\ (n - 2)} \qquad (17)$$

where

x' is the missing value,
X_p', X_v', X_h' are the sums for the indenter, row and column, respectively, in which one value is missing,
n is the number of square elements.

<u>Example 3. Missing value</u>

Let us assume that in Example 1 the value in the 4th row and 1st column is missing. By employing formula (17):

$$x' = \frac{5\ (-5 + 52 + 85) - (2 \times 229)}{4 \times 3} = 17$$

The estimation is identical with the measured value. Of course this is not always the case. E.g. if the value of the 4th row and 2nd column is missing:

$$x' = \frac{5\ (19 + 53 + 17) - (2 \times 253)}{4 \times 3} = -5$$

while the measured value was - 7.

Or if the value in the 5th row and 4th column (0) is missing, the estimation gives x' = - 5.

The value of measurement uncertainty is not much influenced by a missing value. Take into consideration that the estimation for uncertainty was found in this case to be ± 7.

3.6. Use of the analysis of variance in the metrology of hardness

Evaluation of Latin Square experiments is only one of the possible applications of the analysis of variance in the metrology of hardness. This was discussed here in detail so as to get acquainted with the method. Numerous other possibilities of the application of the analysis of variance were published in the literature on hardness measurement, especially in Japan.

In the following the main data of these applications are summarized. The referenced detailed descriptions can serve as models at planning experiments.

a) Comparison of the uncertainty and sensitivity of various hardness testing methods [W-2].

Factors: Testing methods
Blocks

b) Stability of hardness testers [Y-17]

Factors: Areas on block surface
Replication
Interaction

c) Fluctuation of hardness testers over a long period of time [Y-4]

Factors: Machines (40)
Blocks (5)

d) Hardness correction values [Y-4]

Factors: Variation of the machine in time
Indenters
Blocks
Hardness testers
Interactions

e) Evaluation of the error of a Brinell hardness standard [S-11]

Factors: Machines (2)
Hardness levels (4)
Indenting velocity (3)
Duration of loading (3)
Test force (3)
Deformation of the indenter (3)
Replication (2)
Magnification (3)
Hardness level - magnification interaction

f) International hardness comparison [Y-20]

Factors: Indenter, P (2)
Replication, R (4)
P x R interaction
Blocks, B (12)
R x B interaction
P x R interaction
R x P x B interaction
Areas on block surface, A (3)
A x B interaction

g) Hardness distribution on block surface [Y-4]

Factors: Radial direction (3 levels)
Circular direction (5 levels)
Repetition

h) Indenter corrections [Y-12]

Factors: Indenters, P
Blocks, B
P x B interaction
Areas on block surface, A

i) Machine-indenter interactions [P-7]

Factors: Indenters
Machines
Interaction

j) Personal differences in measuring Vickers indentations [Y-19] (several evaluation methods)

Factors: Skill of persons
Persons, W
Size of indentation, M
W x M interaction
Direction of measurement, N
W x N interaction

k) Evaluation of different coincidence methods on a microscope [T-5]

Factors: Methods
Observers
Replication
Tip angle of the mark
Observer-angle interaction

l) Form measurement of small spherical surfaces [I-1]

Factors: Balls
Replication
Interaction

4. Uncertainty problems of indenters

The effect of manufacturing tolerances for indenters on measured hardness values was discussed in Chapter 8 of the BIML Publication "Factors influencing hardness measurement" [P-21].

The performance test of the indenter, which was discussed in Chapter 9 of the BIML Publication "Hardness test blocks and indenters" [P-22] establishes a correction value with respect to a standard indenter or to a reference value. The correction value is a function of the hardness level.

It would be desirable to use indenters with geometrical parameters exactly at nominal value (i.e. desirable tolerance ± 0) for standardization purposes. But this is practically impossible. The standards specify tolerances for the geometry of indenters, e.g. for cone angle and radius of the spherical tip in the case of the Rockwell C method.

As shown in [B-5] and [P-21] the geometrical tolerances of the indenter alone may be responsible for hardness measurement errors of ± 0.7 HRC. Even indenters having identical geometrical parameters (within the uncertainty of measuring these parameters) may give hardness values differing by ± 0.3 HRC at high hardness levels. (The determination of the tip radius is less reliable than that of the cone angle. And at high hardness levels just the spherical part of the indenter is decisive in producing the indentation.) Even higher measurement errors may be caused by those characteristics of indenters, for which no specification (tolerance) can be set up, such as surface roughness, or in general surface quality, transition between spherical and conical part of the Rockwell indenter, rotational symmetry, coincidence of rotation axis with indenter axis, etc.

4.1. Collective standard of indenters

An indenter is a delicate instrument (or tool) which may be hurt or broken at any time. Therefore the use of a single indenter as a standard is not a reliable solution. Most institutes established a collective standard of indenters, consisting of several indenters. In the case of the destruction of one element in the collective standard the standard (reference hardness value) can be maintained and the missing element replaced by another indenter.

The collective standard or group standard of indenters consists of 3, 5 or more indenters. The mean of hardness values obtained by each indenter in the group, or some function of the mean, is considered as the reference value of hardness. Accordingly each indenter in the group has a correction value with respect to the reference value.

HORMUTH [H-10] used at first 5 HRC-indenters as a collective standard. The deviations from the reference value as shown in Fig. 10 were determined by 10 measurements with each indenter at each hardness level. Later 7 indenders were taken to define the reference value. From the 7 indenters several groups of 5 were used as collective standards, the deviations of which are shown in Fig. 11. It can be seen that each collective standard is within ± 0.1 HRC from the reference value. Taking

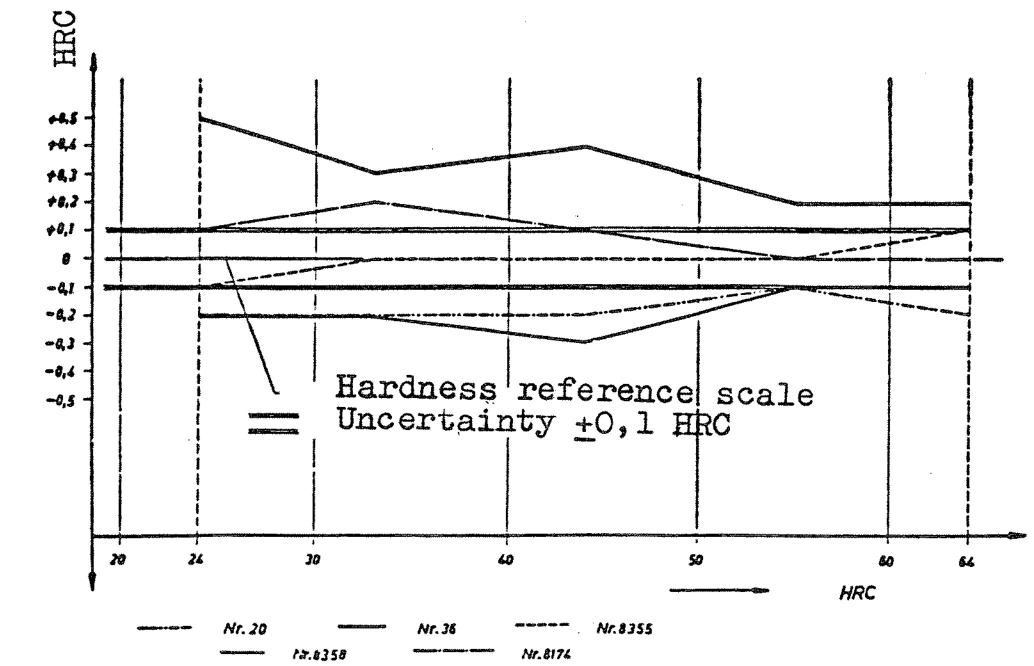

Fig. 10. Deviations of five HRC-indenters from the reference value established as their mean.

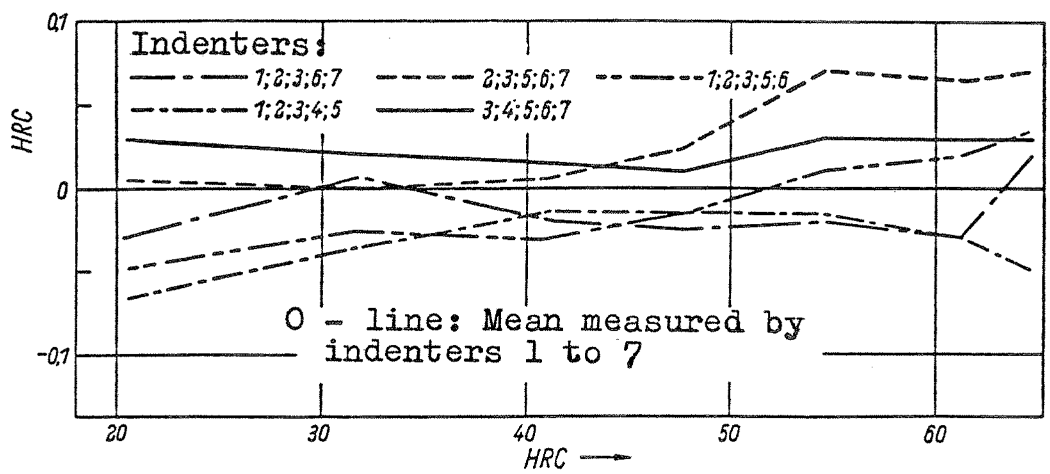

Fig. 11. Deviations of various HRC-indenter collective standards of 5 elements each, from the reference value established as the mean of seven indenters.

the uncertainties of measurement on hardness standard machines into consideration, these deviations can be statistically regarded as zero.

The HRC-indenter collective standard developed by ČUTKA [C-1, C-7] consists of three elements. Ten indentations were made by each indenter on 11 blocks. This measurement series was performed three times. After the elapse of six months the whole procedure was repeated again. The indenter corrections established as the mean of all measurements were represented by a regression line of the first order, in the following form:

$$H = H_i \cdot k + k_o$$

where

H is the hardness in terms of the reference scale,
H_i:the hardness measured by the respective indenter.
k, k_o: constants.

E.g. for one of the elements of the collective standard (No. 3349): $k = 1.0019$, $k_o = - 0.24$.

If this indenter is used for HRA measurements, no correction is necessary ($k = 1$, $k_o = 0$).

A measurement series performed three years later gave values within ± 0.05 HRC from earlier regression lines.

In a collective standard, indenters from the same manufacturer should be used. The manufacturing process itself may produce certain "systematic differences" with respect to other types of indenters. In Fig. 12 the deviations of four indenters of the same manufacturer from values obtained by an indenter of another manufacturer are shown [P-20].

The character of the four curves is apparently similar. The group formed of these 4 + 1 indenters is not convenient as a collective standard. When four indenters of the same manufacturer were added to the original standard indenter (called + 1 above), correction curves shown in Fig. 13 were obtained. For each indenter two curves determined in two different institutes are shown. These five indenters form the HRC-indenter collective standard of Hungary.

In Japan a group of ten indenters forms the collective standard [Y-6]. The absolute deviation of hardness values measured by these indenters from those of the ideal indenter were found to be less than 0.22 HRC.

The deviations of a collective standard of HV-indenters consisting of four elements from the reference value obtained as their mean are shown in Fig. 14 [H-8]. Considering uncertainties from other error sources of Vickers measurements, HV indenter corrections can often be neglected. The Vickers indenter form is more compatible with the crystal characteristics of diamond. This helps to approximate ideal geometry for Vickers indenters much better than in the Rockwell C case.

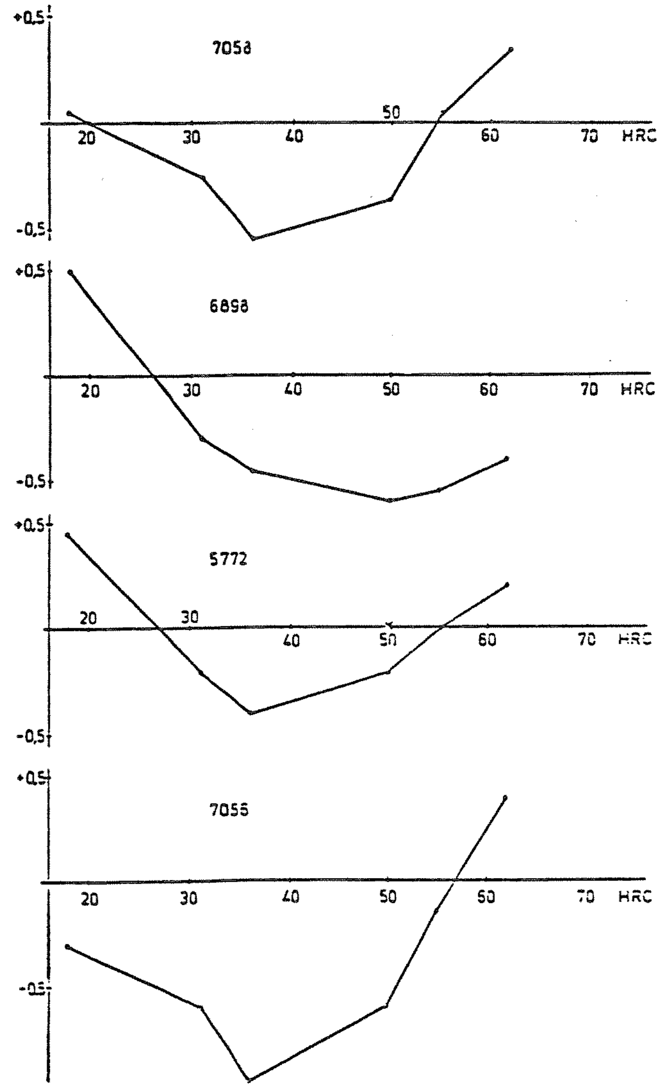

Fig. 12. Deviations of values measured by four different indenters of the same manufacturer from values measured by an indenter of another manufacturer.

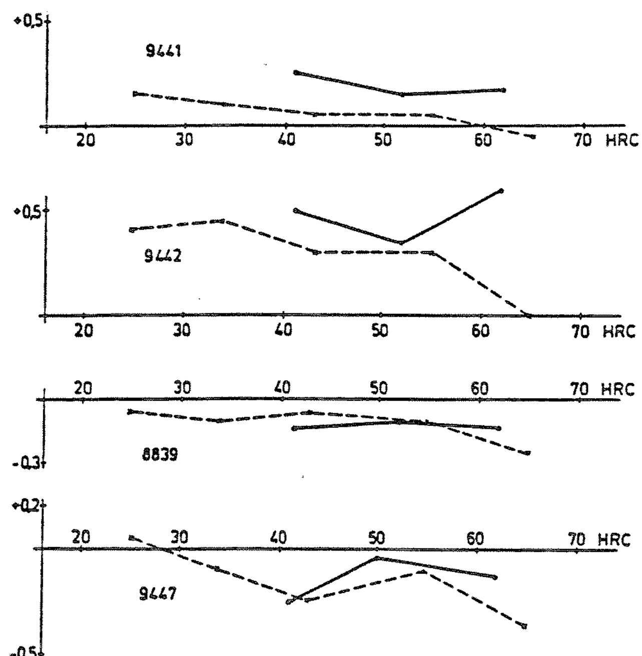

Fig. 13. Deviations of values measured by four different indenters from those measured by another indenter. All five of the same manufacturer.

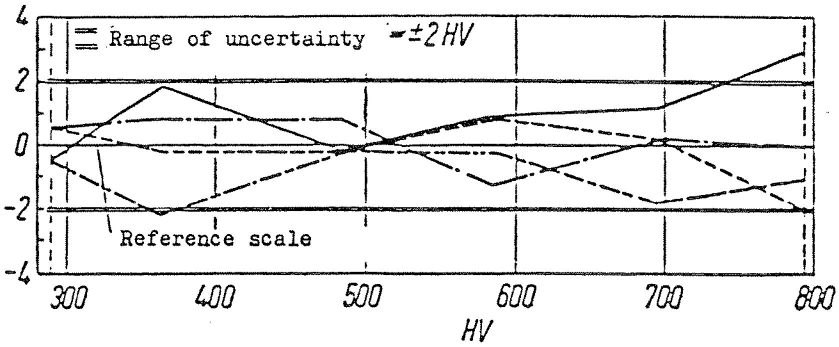

Fig. 14. Deviations of four HV-indenters forming a collective standard from the reference value.

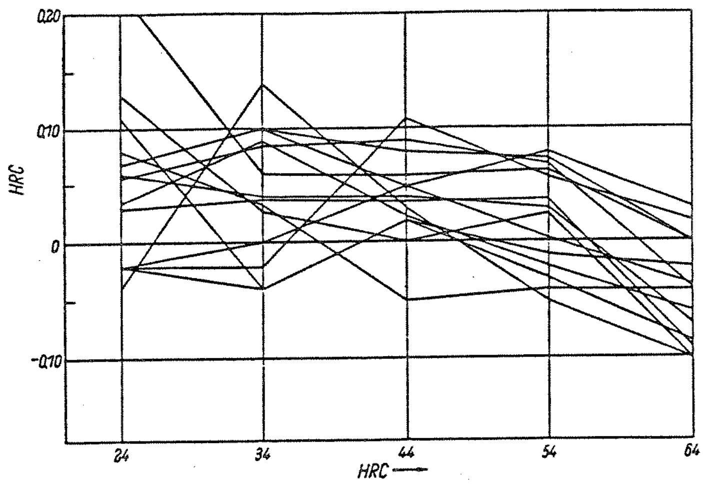

Fig. 15. Repeated determination of the corrections of an indenter. 14 repetitions during 8 years.

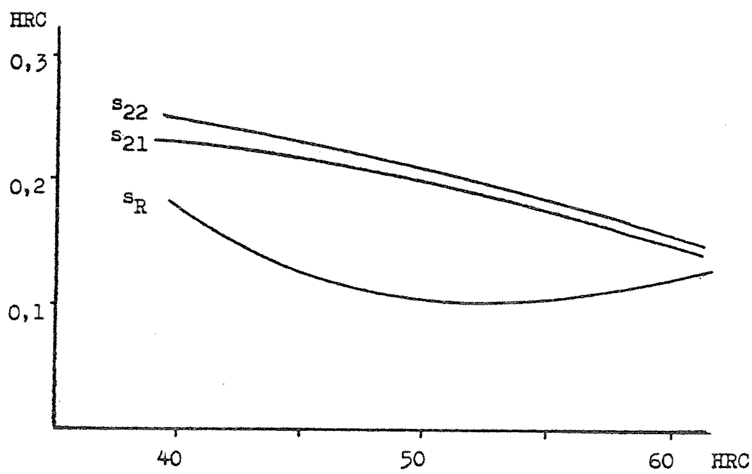

Fig. 16. Estimations for the uncertainty of indenter correction curves.

4.2. Uncertainty of the indenter correction

Indenter corrections given in the form of broken line diagrams (Fig. 10) or as regression lines are not exactly defined values, they have an uncertainty. E.g. one of the lines given in Fig. 10 was determined by 14 repetitions of the measurement series. During a period of 8 years the curves shown on Fig. 15 were obtained. In this form the results cannot be used. A single line shall be used, whereby the existence of an uncertainty of the determination of the line should be borne in mind.

In other words the problem is to find the random error of the determination of a systematic error (correction).

Uncertainty can be estimated by the following methods:

a) According to point 2.3 the uncertainty of the hardness value of a block, as determined by n_1 measurements is

$$s_{12} = [u_1^2 + (s_2^2 + u_3^2)n_1]^{1/2}$$

Indenter correction is defined as a difference with respect to the mean determined by five indenters. The uncertainty of the mean is found to be

$$\bar{s}_{12} = \frac{s_{12}}{\sqrt{5}}$$

The uncertainty of the correction, s_{21}, as the difference of these two values is found to be

$$s_{21} = (s_{12}^2 + \bar{s}_{12}^2)^{1/2} = 1.1\ s_{12}$$

For $n_1 = 5$ measurements on blocks with a reduced range of uncertainty ($R = 0.01\ e$), calculated s_{21} values are plotted in Fig. 16.

b) If correction values of an indenter are repeatedly determined on different standard machines, also indenter-machine interaction should be taken into consideration. It will be shown in point 4.4 that this interaction is approximately equal to s_2. Accordingly the uncertainty s_{22} of the indenter correction, if determined on five machines, is found to be

$$s_{22} = \left(s_{21}^2 + \frac{s_2^2}{5}\right)^{1/2}$$

Calculated values are shown in Fig. 16.

c) The uncertainty of the correction curves can be determined also from experimental values [P-10].

In the test series 5 indenters, 7 blocks, and 5 hardness standard machines were used. Five HRC indentations were made in each combination of the above-mentioned three factors (5 x 7 x 5 = 175), by a single observer. Thus the total test series consisted of 875 indentations.

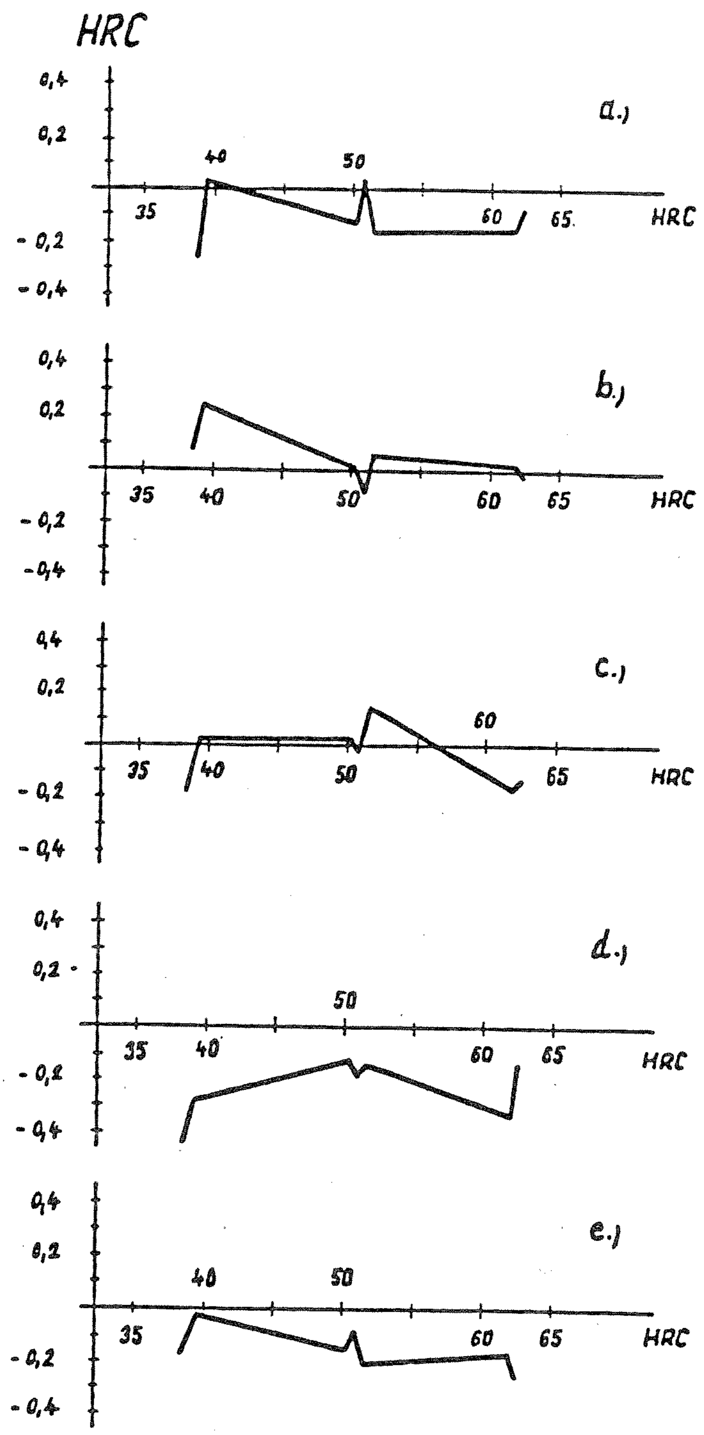

Fig. 17. Corrections curves of the same indenter with respect to the group mean, as determined on five different machines.

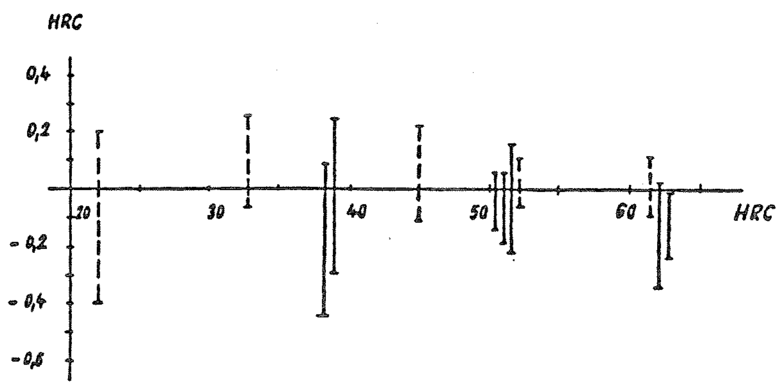

Fig. 18. Ranges covered by correction curves in Fig. 17.

(The results of another experiment with dotted lines)

Mean values of the five measurements made with the same indenter on the same machine and on the same block supplied one point of the correction curve. In Fig. 17 correction curves of one of the indenters as determined on five machines are shown.

Differences between the five curves indicate that none of them can be regarded as the one supplying the "correct" correction values. In trying to find the reliability of correction curves of this kind, the following considerations were employed.

Fig. 18 was constructed, where vertical sections indicate the range of points taken over from Fig. 17 a to e. In Fig. 18 the ranges of values for the same indenter, from a former measurement [P-4] are also indicated by dotted lines, for purposes of comparison and checking. There was no significant difference between ranges obtained for the correction curve in the two experiments. By using formulae of mathematical statistics, an estimation of standard deviations can be obtained from the range of values.

s_R values corresponding to ranges shown in Fig. 18, as well as those for similar diagrams constructed for the other four indenters were calculated and plotted in function of nominal hardness. From five values for each hardness level, the second order regression curve (s_R) shown in Fig. 16 was obtained.

The three curves shown in Fig. 16 give three estimations of the uncertainty of the indenter correction determined by different methods described above. The small difference of curves s_{21} and s_{22} shows that the machine-indenter interaction (use of various machines to establish the correction) has a relatively small effect on uncertainty.

YANO et al. [Y-12] employed a multiple regression analysis to find the "ideal" indenter. They found that after corrections employed to the standard indenter, the ideal form can be found within the following standard deviation values: 0.08 HRC, 0.08 HRA, 0.10 HR30N. The range of errors of the primary calibration of indenters, at the 99 % confidence level: ± 0.12 HRC, ± 0.08 HRA and ± 0.14 HR30N. The same for secondary indenters: ± 0.16 HRC, ± 0.10 HRA and ± 0.20 HR30N. Indenter corrections which are smaller than 0.04 HRC, 0.03 HRA or 0.05 HR30N are neglected as not significant.

4.3. Indenter-block interaction

In an international HRC comparison YAMAMOTO and YANO [Y-20] observed considerable fluctuations in machine differences if the hardness of two blocks differs slightly, say 1 or 2 HRC. The reason for this cannot be clarified strictly, but it is possible that the fluctuations originate from the errors of indentation depth measurement. The above mentioned authors did not think that indentation loads or form errors of indenters may have such an effect in the case of small hardness differences.

PETIK [P-10] confirmed these observations. In the experiment described in point 4.2, two blocks each for the hardness levels of approximately 38 HRC and 62 HRC, and three for the level of 51 HRC were used. The results given in Fig. 17 clearly show the fluctuations of the correction curves at small hardness differences.

In a personal communication HILD expressed the opinion that these fluctuations may simply originate from measurement uncertainties.

YANO et al. [Y-12] made a series of experiments in which four blocks were used at each hardness level. On the surface of each block three sections were marked. In each of them, indentations were made with 20 or 30 indenters. The analysis of variance showed a significant interaction between blocks and indenters. This interaction may be responsible for the previously mentioned fluctuations on the correction curve.

This question deserves further research work.

4.4. Indenter-machine interaction

When comparing different hardness testers it was found that part of the difference was due to indenters, another part to machines. Therefore comparisons are often made in such a way that the same (travelling) indenter is used in both machines. A comparison of this kind establishes the difference of the two machines. By measurements made on the same block, on the same machine with the travelling indenter on the one hand, and by the own indenter of the machine on the other, the difference of the two indenters can be established. By employing this method, the correction value of the hardness tester (machine and indenter) is supposed to be the sum of the corrections for the machine and for the indenter. But the two corrections are additive only if the machine and indenter effects are independent, i.e. there is no interaction between the two.

To verify the hypothesis of independent effects, the measurement values of the experiment described in point 4.2, consisting of 5 x 7 x 5 = 175 combinations of indenters, blocks and machines were evaluated [P-7].

The 125 values measured by a single person on each block were processed by the analysis of variance. As an example, Table 8 shows the analysis of values obtained on the block, which had an overall mean of 50.92 HRC.

Table 8

Analysis of variance for a machine-indenter interaction test

Source	Sum of squares, SS	Degrees of freedom, f	Mean square MS = SS/f	Fischer test value F
Indenters	1.9245	4	s^2_{IV} = 0.4811	s^2_{IV}/s^2_I = 26.5
Machines	1.0069	4	s^2_{III} = 0.2517	s^2_{III}/s^2_I = 13.9
Interaction	0.8493	16	s^2_{II} = 0.0531	s^2_{II}/s^2_I = 2.93
Residual	1.8150	100	s^2_I = 0.0182	
Overall	5.5957	124		

s^2_I and s^2_{II} supply estimations for the following values:

$$s^2_I \mathrel{\hat{=}} \sigma^2$$

$$s^2_{II} \mathrel{\hat{=}} 5\ \sigma^2_\alpha + \sigma^2$$

where

σ_α represents the machine indenter interaction

and σ is characteristic of the measuring equipment and method.

The critical F-values for the statistical test are

$$F_{cr}\ (16;100)_{0.95} = 1.75$$

$$F_{cr}\ (16;100)_{0.99} = 2.19$$

F-values for interaction, considering also the tables for the other six blocks, were found to be in the range from 2.6 to 7.6, consequently the interaction was always significant at the 0.99 confidence level.

This series of measurements proved the existence of an interaction between the standard machine and the standard indenter, accordingly correction values for the machine and the indenter cannot be simply added. Machine and indenter should be considered as a single, inseparable unit. In the contrary case, σ_α should be taken into consideration at specifying uncertainty of the standard equipment. σ_α values were found to be of the same order of magnitude as the measurement uncertainty σ. Fig. 19 shows σ and σ_α values found at different hardness levels. In

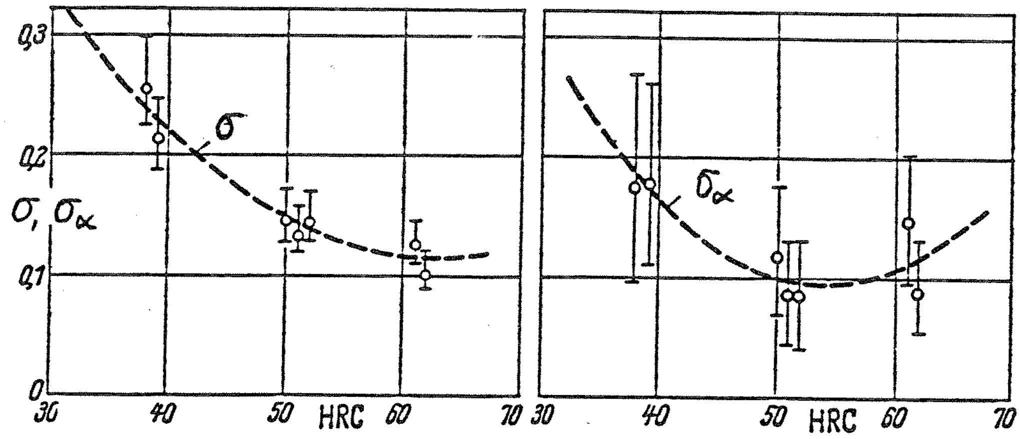

Fig. 19. Measurement uncertainty (σ) and machine-indenter interaction (σ_α) as determined on 5 machines with 5 indenters by a single person.

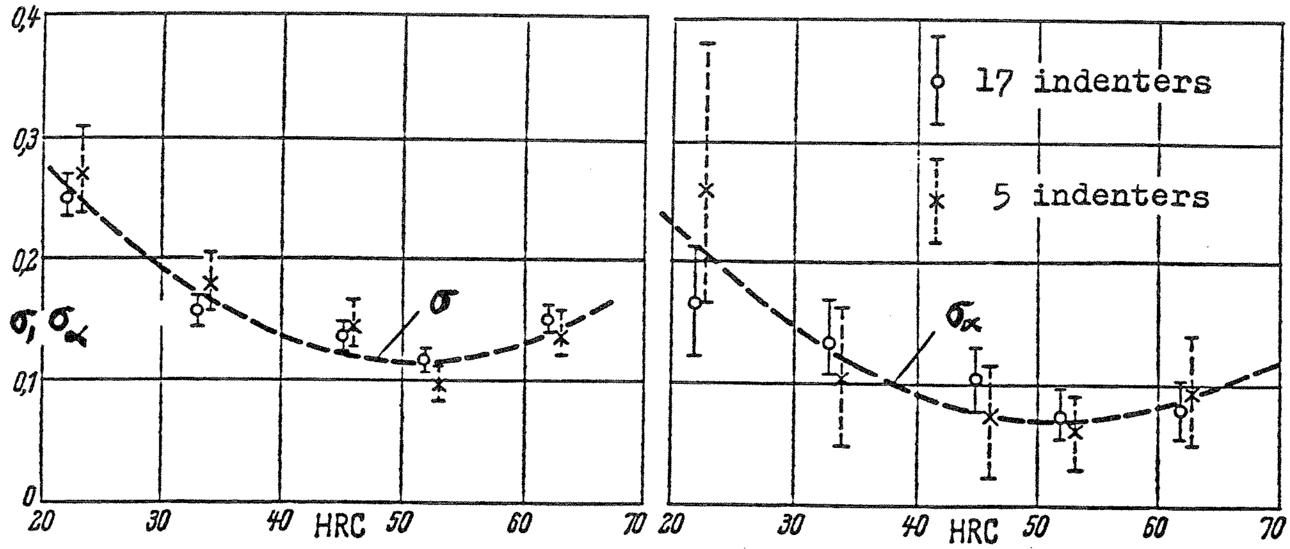

Fig. 20. Measurement uncertainty (σ) and machine-indenter interaction (σ_α) as determined on 5 machines with 17 (respectively 5) indenters, by 5 persons.

$$(1-\alpha)100 = 95\ \%$$

are also indicated, calculated at the left side of Fig. 19 by the formula

$$\frac{f}{\chi^2_{1-(\alpha/2);f}}\ s_I^2 \le \sigma^2 \le \frac{f}{\chi^2_{\alpha/2;f}}\ s_I^2$$

where

$f = 100,\qquad \chi^2_{0.975;100} = 129.6$

$\chi^2_{0.025;100} = 74.2$

For the standard deviation σ_α representing interaction (right side of Fig. 19) the confidence range is defined as

$$\frac{1}{5}\left[\frac{s_{II}^2}{F_{1-(\alpha/2)}(f_1;f_2)} - s_I^2\right] \le \sigma_\alpha^2 \le \frac{1}{5}\left[s_{II}^2\ F_{1-(\alpha/2)}\ (f_1;f_2) - s_I^2\right]$$

where

$f_1 = f_2 = 100$ and $F_{0.975}\ (100;100) = 1.94.$

The broken lines drawn in Fig.19 are calculated regression curves of the second order.

To check the values obtained in this interaction experiment, the results of an earlier international comparison [P-4], originally organized with other aims, were also evaluated statistically. In that comparison five persons were working on five standard machines, each person having a set of five blocks of different hardness levels. 17 indenters were used alternately on each machine and each block, producing five indentations on each. The results of the comparison were already discussed in [P-28], here only the results of the statistical evaluation of the machine-indenter interaction are given.

The number of degrees of freedom:

Source	f
Between indenters	16
Between machines	4
Interaction	64
Residual	340
Overall	424

The concept of variance between machines is not identical here with that in the previously discussed experiment. Here it would be better to speak of "variance between machines, including measuring person and his set of blocks".

Results are shown in Fig 20 (values marked by small cercles). Values

Results are shown in Fig 20 (values marked by small cercles). Values used for establishing confidence ranges:

$$f_1 = 64, \; f_2 = 340, \; F_{0.975}\,(64;340) = 1.44.$$

All interaction values were found to be significant.

To have a better possibility of comparison with the results of the analysis of variance of the experiment represented by Table 8, the results were evaluated also in such a way that only 5 indenters from among the 17 were taken into consideration. Results are shown in Fig. 20 by the points marked by an x. In this case too all interaction values were found to be significant.

The values shown in Fig. 19 and 20, respectively, do not differ significantly.

The existence of a machine-indenter interaction, as evidenced by the experiments described here, were confirmed by ECKARDT [E-4]. The reason necessitating this series of experiments was the following.

In a collective standard of five HRC indenters all correction values were within a band which is about 0.3 HRC large. The corrections for the same indenters were determined also on another standard machine. The band in which new correction values were included increased to 0.6-0.7 HRC. Within this band the relative positions of the individual correction curves also changed. The series of planned experiments, with analysis of variance gave similar conclusions as the previously discussed experiments. The most surprising result of this series was that two secondary standard machines were measuring about 0.2 HRC below the national standard with one indenter, and about 0.1 HRC above the national standard when using another indenter. This difference can be attributed to the machine-indenter interaction.

5. International hardness comparisons

5.1. Methods of comparison employed

In international hardness comparisons the standard equipments of two or more countries are compared. Standard equipment serves to maintain the national hardness reference scales of the respective country. Thus the international hardness comparison determines the difference between different national hardness reference scales.

The standard equipment (machine, microscope) is stationary, cannot be transported. The only element of the standard which can be transported is the indenter.

The connection between stationary hardness standard equipments is ensured by the hardness reference blocks, serving as transfer standards.

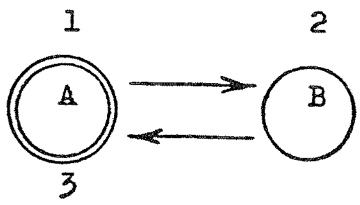

Fig. 21. Scheme of a bilateral comparison.

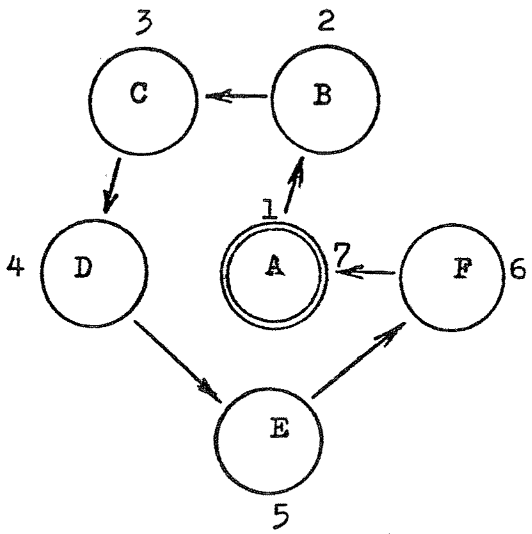

Fig. 22. Scheme of a simple multi-lateral comparison.

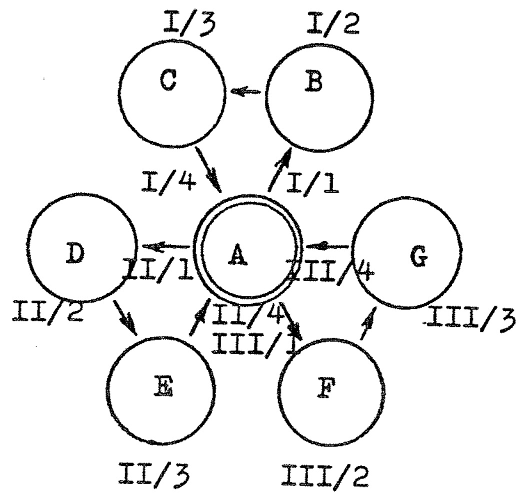

Fig. 23. Scheme of a multilateral comparison with subgroups.

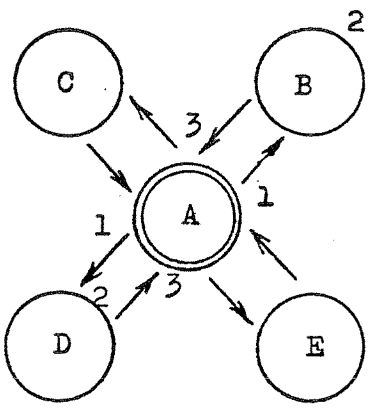

Fig. 24. Scheme of a multilateral comparison with different sets of blocks in each relation.

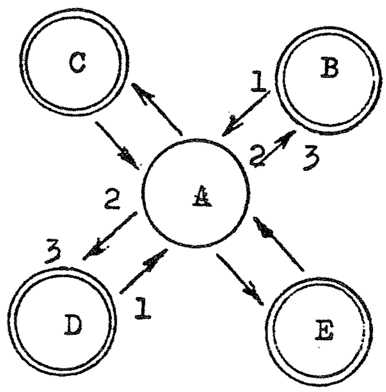

Fig. 25. Another variant of the scheme shown in Fig. 24.

In international hardness comparisons one of the laboratories sends to another laboratory:

- a set of blocks, or
- a set of blocks and indenters, which are
- in certain cases accompanied by a measuring person (especially useful in the case of the Vickers and Brinell method).

The simplest case of a bilateral comparison is shown in Fig. 21. The circles represent laboratories, the double circle indicates the laboratory sending the blocks. It is usual to measure the block before sending and after the return. The numbers in the figure indicate the order of measurements.

There are several possible methods for organizing multilateral comparisons. The less expensive but most time consuming method is represented in Fig. 22. In this case a single set of blocks is sent around the six participating laboratories.

The duration of the comparison can be reduced by forming subgroups of laboratories using different sets of blocks (Fig. 23).

A further step is to send different sets of blocks to each participant (Fig. 24), or each participant sends his blocks to the organizing laboratory (Fig. 25). This method is practically a multiple bilateral comparison.

Other comparison schemes are also possible.

Before starting the comparison, all the details must be agreed and set down in written instructions. These instructions must indicate, among others

- the number of indentations to be made,
- the place of indentations on the blocks,
- the time schedule of making the indentations,
- specification of the test cycle,
- cleaning the blocks before the tests,
- preparation of the blocks before sending them back,
- unified tables for the presentation of measured data,
- statistical methods to be employed for evaluation.

A remark concerning statistics is useful here. Statistical evaluation has become on inseparable part of metrological work. Methods should be used, however, cleverly, after careful considerations, never mechanically. In some cases statistics can be used (or misused) to prove some desirable conclusion, or to reject the same, if desired so. In a comparison of two measuring instruments (standards) the difference of results obtained on both is determined. The question, whether this difference is significant, is answered by using statistical methods. There are cases when the metrologist is interested to prove that the difference is significant. ("As a result of our work we were able to prove that our partner measures differently".) In other cases the metrologist is personally interested to have a not significant difference. ("The observed difference between our standard and the international standard is not significant, consequently our standard is just as good".)

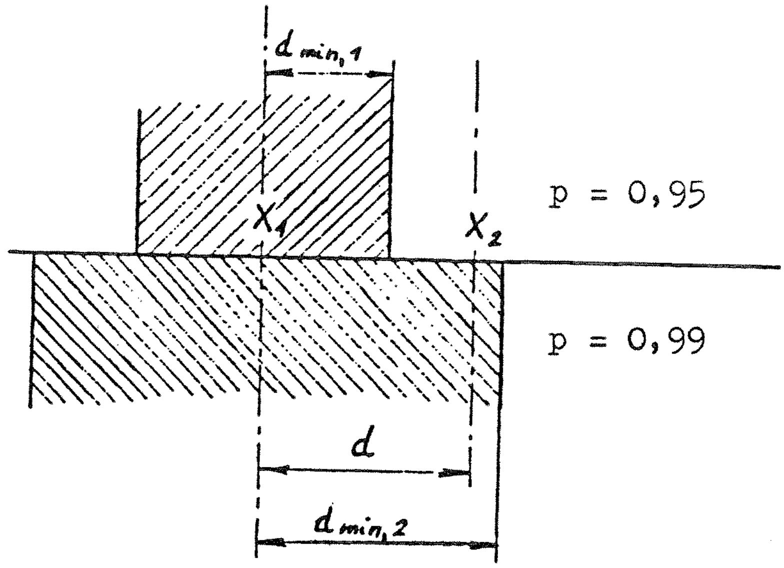

Fig. 26. Schematic illustration of the significance of measured difference $d = X_2 - X_1$

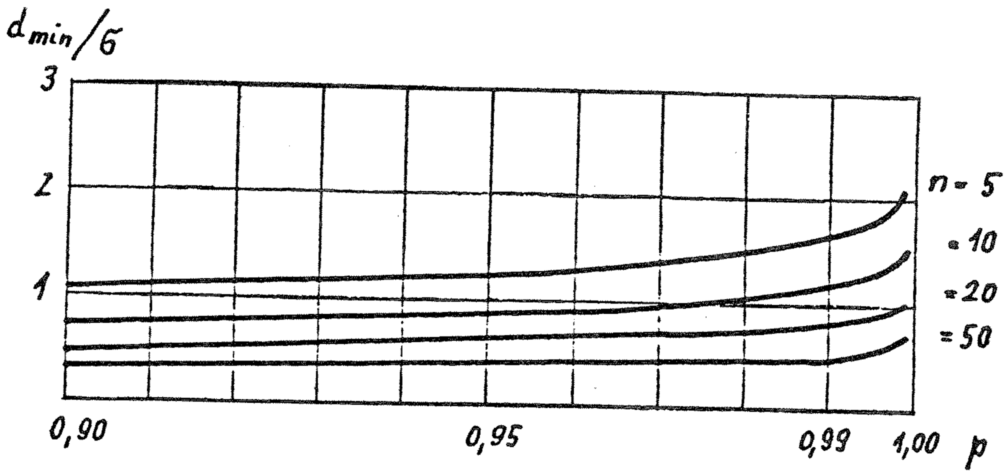

Fig. 27. Minimum relative differences at which the zero hypothesis can be rejected.

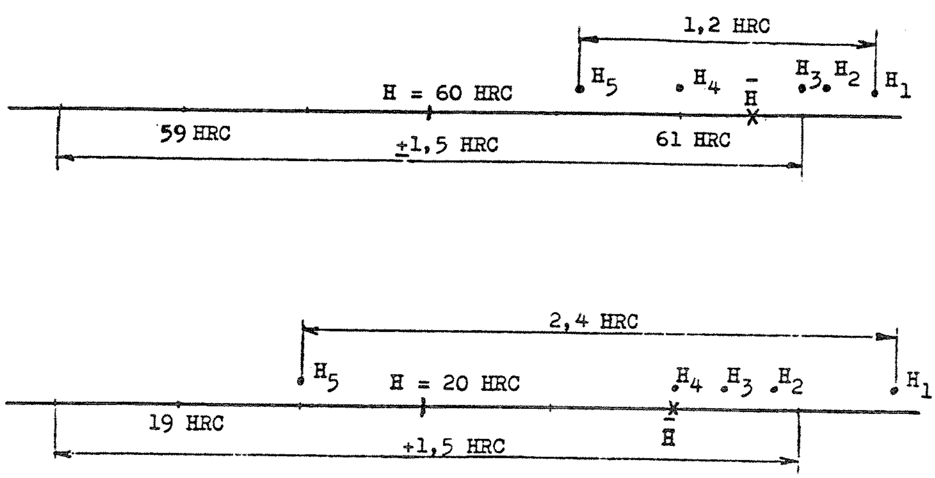

Fig. 28. Tolerances specified in ISO 716 for a commercial hardness tester.

The following considerations show how such contradictory arguments can be justified by employing statistical methods [P-9].

Let X_1 and X_2 designate the mean of n measurements made by two measuring instruments (standards). Their difference is

$$d = X_2 - X_1$$

To determine whether d is significant the Student-test is employed. By this test a d_{min} value can be obtained. If

	$d < d_{min}$,	d is not significant,
if	$d > d_{min}$,	d is significant.

In the upper part of Fig. 26, $d_{min,1}$ represents the limits for $p = 0.95$ confidence level. In the lower part $d_{min,2}$ the same for $p = 0.99$ confidence level. It is evident that in the case represented by the figure, the difference d is significant at $p = 0.95$. By simply raising the confidence level to $p = 0.99$, difference d is not significant any more. In a better formulation: At $p = 0.99$ we cannot reject the hypothesis that $X_1 = X_2$.

To avoid similar ambiguities, all details of statistical evaluation, especially confidence levels should be fixed in advance.

Fig. 27 shows d_{min}/σ values within which the zero hypothesis cannot be rejected, at various confidence levels (p) and repetitions (n).

5.2. Published comparison results

The results of numerous international hardness comparisons were published. Others were circulated in the restricted circle of the participants.

Table 9 summarizes in a chronological order the accessible publications.

Comparison results should not be used as correction values without further considerations and consultations. Some of the national hardness reference scales, the relative position of various scales has changed with time, as it can be seen from an evaluation of comparisons. Some intentional changes are discussed in point 5.3. Other changes have unknown or not published reasons.

If the knowledge of the differences of two national hardness reference scales is necessary for practical reasons (delivery of certain products), only recent comparison results should be taken into consideration, after consultations between partners.

Table 9

Published data on international hardness comparisons

Reference	Year of publication	Test methods	Participating countries	Essential features of the comparisons
M-11	1957	HRC	AT,DD,DE,F,GB,JP,SE,US	
M-4	1965	HRC	GB,JP	
F-1	1966	HRC	AT,DE	6 blocks, own and travelling indenter
H-10	1966	HRC HV 30	BG,CS,DD,HU,JP,RO,SU CS,DD,PL,RO,SU	
P-1	1966	HRC	DD,DE,HU,RO,SU	
R-2	1966	HB	DD,RO	
Y-20 Y-5	1966	HRC	AT,CA,CN,DD,DE,GB,JP SU,US	12 blocks (of 6 hardness levels) and 1 indenter sent to each participant. Local indenter also used. Three indentations on each block and by each indenter. Second repetition after a week, third after another week.
P-2	1967	HRC	DD,HU	
P-3	1967	HRC	AT,CS,DD,HU,SU	
S-8	1967	HRC	AT,DD,DE,FR,GB,JP,PL SE,SU,US	

Table 9 (continued)

P-4	1970	HRC,HV,HB	BG,CS,DD,HU,PL	Relative differences of 16 indenters determined independently by 5 persons on four machines, using 5 blocks each, 5 indentations in each combination - Measurement of 3 x 15 HV indentations by five persons on three microscopes, with repetitions on the following two days. The same for 2x18 HB indentations on two microscopes.
M-9	1972	HRC,HV	AT,AU,DD,EG,JP,PL,US	
P-6	1972	HRC	GB,HU	6 blocks calibrated alternately in two laboratories, four times in each, during a period of 2 1/2 years. Two interchanges made also with travelling indenter.
I-2	1976	HRC,HV	AT,DD,GB,HU,JP	HRC and HV indentations made on same blocks. Three indentations per block, two repetitions.
P-12	1978	HRC,HV	BG,CS,DD,HU,PL,RO,SU	7 machines compared by the same person using 6 blocks, making 8 indentations on each. Repetition with travelling indenter. The same four years later using 4 blocks, 7 indentations on each, using 2 local and 2 travelling indenters.
G-2	1980	HRC, HV 30	FR, HU	3 blocks, local and travelling indenter 2 blocks, readings on microscopes of both institutes
T-4	1980	HRC	CN,DD,JP	
W-23	1982	HRC	DE,GB,JP,US	
B-10	1983	HRC HRC	IT and 5 countries AU,CH,CS,DD,GB,HU,IT PL, SU, YU	Exchange of blocks. Without travelling blocks. Indentations were made on HRC machines with both HRC and HV indenters.
C-6	1984	HV 30	CS and 12 countries	
O-3	1984	HRC, HV 30	AU,AT,BG,CS,DD,DE,GB, HU,IT,PL,RO,SU,US	5 blocks, 5 indentations HRC (10 indentations HV) on each, repeated on two following days.

Table 9 (continued)

Y-18	1984	HRC	CN,GB,IT,JP,US	
C-7	1985	HRC	CS and 14 countries	
B-14	1986	HRA,HRB,HRC HRN,HRT,HV	CN,IT	Comparison with own indenter and travelling indenter. Repetition after one year.
K-21	1986	HRC,HV 30	AT,CS,DD,HU,PL	
B-16	1988	HRC	JP,US	
C-9	1988	HRN	CS,DD,GB,PL,US	
C-11	1988	HRC	BG,CS,DD,HU,PL,RO,SU	3 blocks
		HV 30	BG,CS,DD,HU,RO,SU	3 blocks
		HRC	AT,AU,BG,CS,DE,DD,FI GB,HU,IT,PL,RO,SU,US	On 5 blocks 3 x 5 indentations
		HV 30	AT,AU,BG,CS,DE,DD,FI GB,HU,PL,RO,SU,US	On 5 blocks 3 x 10 indentations
H-20	1990	HRB,HRC	DE,DK,FR,GB,IT	3 blocks in 7 laboratories, 5 indentations each with local and travelling indenter. Second stage: The same in 3 laboratories with 8 blocks.

Country codes: AU - Australia, AT - Austria, BG - Bulgaria, CA - Canada, CN - China, CS - Czechoslovakia, DD - German Democratic Republic, DE - Federal Republic of Germany, DK - Denmark, FI - Finland, FR - France, GB - United Kingdom, HU - Hungary, IT - Italy, JP - Japan, PL - Poland, RO - Romania, SE - Sweden, SU - Soviet Union, US - United States of America, YU - Yugoslavia.

5.3. Redefinition of some national hardness reference scales

As a result of international hardness comparisons, or on other reasons, in some cases a redefinition (shift) of a national hardness reference scale may become necessary. Three known cases deserve mention here.

a) HRC hardness testers produced in the Soviet Union in the 1930's were checked by using imported blocks calibrated by the Wilson Company in the USA [S-8]. HRC values measured in the Soviet industry corresponded accordingly to the Wilson Reference Scale of that time. The use of this HRC scale in industry was maintained until 1980. After the second World War Soviet Metrological Institutes set up deadweight Rockwell standard machines which were compared with standards of other states.In this way the values of the so-called All-Union Standard were conform with other national standards, but differed from hardness values measured in the Soviet Industry. This anomaly was removed in 1979 by publishing the State Standard Specification GOST 8064-79 establishing the legal status of the All-Union Rockwell standard machine and of the hierarchy scheme of Rockwell hardness measurements performed in the Soviet Union. The Standard contains a conversion table between the traditional industrial Rockwell values (denoted by HRC) and the Rockwell values of the standard equipment (denoted by HRC_e). Some characteristic points of the conversion table:

HRC_e	HRC
20.0	17.8
30.0	28.1
40.0	38.4
50.0	48.7
60.0	59.0
65.0	64.1

GOST 8064 specifies also some practical measures to ensure the transition to hardness values conform to international practice [K-14]: Hardness testers used in industry are to be modified; hardness values specified in drawings or other technical documents must be given in HRC_e values; blocks calibrated in HRC values can be used with conversion tables during the two years of validity of their calibrated value.

b) In the United Kingdom, in order to disseminate to industries a standard HRC scale with a level that is commercially viable whilst retaining the stability of the national reference scale, a small corrections was applied to the hardness values assigned to each hardness block calibrated on the NPL deadweight machine [W-23]. The corrections, of less than 1 HRC unit, have been evaluated from a knowledge of the effect of design features in commercial machines which tend to reduce the hardness values, together with an allowance for the slightly harder values obtained when using shorter dwell times of only a few seconds compared with a standardizing time of 30 seconds.

The effect of these technical developments is that HRC values with traceability to NPL in 1982 were about 1 HRC unit lower than they were in the 1970's. This is in agreement with practice in other major industrial countries.

c) The effect of dwell time and of other test conditions on hardness reference scales was examined also by the Technical Committee ISO/TC 164. The changes introduced in the new series of ISO International Standards on hardness testing, elaborated in the 1980's, may have an effect on national hardness reference scales.

Earlier loading cycles employed in industrial hardness testing and at calibrating hardness test blocks were different. In industry the hardness test should be performed as quickly as possible, of course without excessive loss of precision. For block calibration the metrology laboratory has more time, consequently load rise time and dwell times were specified so that the penetration process be completely stabilized, to ensure maximum precision and accuracy.

During the elaboration of the revised International Standards on hardness testing the practical opinion prevailed and was fixed as specification for the loading cycle: The calibration of the blocks should be performed in a loading cycle which is identical with that employed in hardness tests in industry.

What are the possible effects of this decision on existing hardness reference scales [P-26]?

For the calibration of Rockwell blocks we are specially concerned with the following elements of the loading cycle:

- t_o, time of application of the preliminary test force
- t_i, load rise time from preliminary test force to total test force
- t_m, time of application of the total test force.

The values given in the old and new ISO specifications for the calibration of standardized blocks for Rockwell hardness testing machines are shown in Table 10.

Table 10

Values specified in the Standards

	ISO/R 674 (1968)	ISO 674 (1988)
t_o	10 to 20 s	1 to 10 s
t_i	-	2 to 8 s
t_m	30 to 35 s	3 to 5 s

Time of application of the preliminary test force (t_o)

The former practice was to maintain the preliminary force for 10 to 20 s, because approximately 10 s was necessary to stabilize the position of the indenter. The new prescription is 1 to 10 s. DAMBACHER found a difference of 0.3 to 0.4 HRC in the final hardness value of a block of 20 HRC, depending on whether the preliminary test force was maintained for 1 s or 10 s. This tolerance range for time, in the unstabilized part of the indentation cycle is very wide for calibration purposes.

Load rise time (t_i)

The difference of hardness values obtained at the two ends of the range specified for load rise time (2 and 8 s, respectively) is 0.2 to 0.3 HRC at 65 HRC, and 0.2 HRC at 20 HRC.

Time of application of the total test force (t_m)

The value specified earlier was 30 to 35 s, and has now become 3 to 5 s.

Shorter application times mean higher hardness values. This effect is greater on materials of low hardness. Stabilization of penetration is not to be expected before 30 s. Table 11 shows hardness values obtained with different application times.

Table 11

Hardness values (HRC) obtained with a duration of

30 s	5 s	3s
20.0	20.6 to 20.7	20.7 to 20.9
60.0	60.2	60.3

This means that the national hardness reference scales should be corrected by up to 0.9 HRC at the low end and by 0.3 HRC at the high end of the HRC-scale. Even the limiting values of the newly specified range (3 to 5 s) may result in a hardness difference of

0.1 to 0.2 HRC at 20 HRC
0.1 HRC at 60 HRC

Table 12 shows a similar effect in the case of the HRB scale.

Table 12

Hardness values (HRB) obtained with a duration of

15 s	5 s	3 s
55.0	56.1	56.8

Addition of uncertainties resulting from relatively large range of time specifications

Three kinds of uncertainties, originating from the specified ranges of time have been mentioned so far (Table 13).

The resulting uncertainty can be considered in two ways:

a) Let us assume, firstly, that time values within the specified ranges will be employed by calibration personnel at random. In this case a standard deviation can be estimated from the sum by quadratic addition, by using the appropriate formula. the three specified time ranges result in an increase in the uncertainty of the hardness standardizing machine by 0.4 HRC.

b) We may assume, alternatively, that a given national standardizing institute will use only a narrow part of the ranges specified in the standard. But these will not necessarily be indentical to those used by other national institutes. The result will be a systematic difference between the calibrated values obtained by different institutes which may be as great as the arithmetic sum indicated in Table 13, namely 0.7 HRC. The international uniformity of hardness measurements is thus impaired.

Table 13

Range of hardness values (in HRC)

resulting from the range of time specified for	at hardness level 20 HRC	60 HRC
t_o	0.3 to 0.4	0.2
t_i	0.2	0.2 to 0.3
t_m	0.1 to 0.2	0.1
Sum by simple addition	0.7	0.6
Sum by quadratic addition	0.4	0.4

6. Uncertainty of industrial hardness testers

The final aim of all metrological work connected with hardness testing is the metrological assurance of industrial hardness testers. A detailed discussion of problems of these hardness testers goes beyond the scope of the present work. Only some questions are mentioned here, with hints and references for those who intend to study this field more in detail.

Several decades ago, the first step towards the unification of hardness measurements was the survey of the state of hardness testers used in various laboratories, the establishement of their precision and relative measurement deviation (accuracy). Extensive comparisons were made, the essential conclusions of which form now the basis of ISO International Standards on the verification of hardness testers (ISO 716-1986, ISO 1079-1989, ISO 146-1984, ISO 156-1982 for the Rockwell, Vickers and Brinell methods). These Standards can be regarded as the condensed outcome of many years of experience.

The metrological behaviour of hardness testers is specified in the standards by two characteristics:

<u>Repeatability</u>, i.e. the range of five consecutive measurements made on a standardized block.

<u>Error</u> i.e. the difference between the mean of the five measurements and the specified hardness of the standardized block.

Fig. 28 illustrates this specification on the example of Rockwell testers at the levels of 60 and 20 HRC respectively. The range specified for the error is identical (± 1.5 HRC) at all hardness levels. The range for repeatability depends on hardness. H is the hardness of the standardized block used. H_1, H_2,...H_5 are the five measured hardness values, $\bar{H}$ the mean of the five.

Specifications given in a standard represent minimum requirements necessary for the use of the hardness tester. In special application more severe requirements, or other specifications better suited for the given task may be employed.

In such cases some of the experience of earlier research work may be useful, such as

HILD [H-4] on the examination of Rockwell testers,
JEGADEN et al. [J-2] and YAMASHIRO et al. [Y-10] on the examination of Super-Rockwell testers,
PILIPTCHUK [P-14] on the theoretical foundations of the fine adjustment of hardness testers,
ROSSOW [R-13] on the examination of Rockwell C testers; with well founded design of experiments and evaluation of results.

Surveys on the state of industrial hardness testing in various countries can be found in the publications enumerated in Table 14.

<u>Table 14</u>

State of hardness testing in given countries

Author	Reference	Country	Year of publication
Meyer	M-11	F.R. Germany	1957
Rossow	R-12	F.R. Germany	1961
Yoshizawa	Y-13	Japan	1961
Rossow	R-13	F.R. Germany	1967
Čutka	C-3	Czechoslovakia	1970
Yano et al	Y-26	Japan	1972
Shinozuka-Matsuura	S-18	Japan	1973
Eckardt-Kersten	E-1	GDR	1974
Shinozuka	S-19	Japan	1975
Dambacher	D-1	F.R. Germany	1978

The real problems facing industrial hardness testing can be better understood by the following considerations, by the help of an imaginary example on the heat treatment of a machine component [B-2].

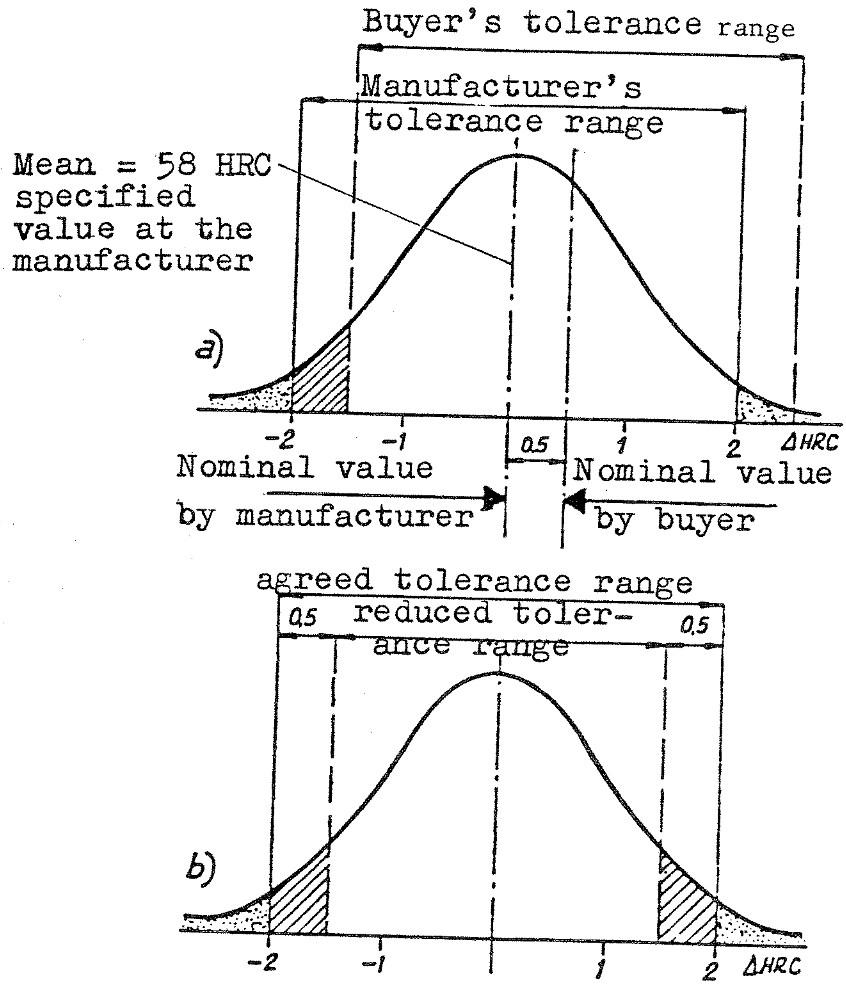

Fig. 29. Frequency distribution of the hardness of heat treated components.

On Fig. 29 the frequency distribution of the hardness of a batch of heat treated components is shown, which is supposed to be Gaussian. The standard deviation of measured hardness values was found to be ± 1 HRC.

The value specified on the drawing of the component is (58 ± 2)HRC. Let us suppose that the mean of measured hardness values corresponds to the prescribed value, namely 58 HRC (what is of course an optimistic assumption). In this case the tolerance range of ± 2 HRC includes 95 % of all produced pieces, 5 % (area shaded by points) is rejected by the manufacturer. The pieces are delivered to another company whose hardness tester has a systematic deviation of 0.5 HRC with respect to that of the manufacturer. This means the rejection of another 5 % of the delivered pieces (shaded area under the curve) when control tests are made.

So far the possible errors of hardness testers were not fully taken into consideration. According to ISO 716 the error of the hardness tester may be ± 1.5 HRC with respect to the block of the verification officer. Thus in our example, in an extreme case, the systematic difference between measurements made by the manufacturer and the receiver may amount to 3 HRC.

This may result in disputes which can be resolved only by direct hardness comparisons between the two partners.

Another problem is the random error of the testers. ISO 716 specifies a repeatibility of 1.2 HRC at 58 HRC, i.e. in the case of five repetitions of the hardness measurement, the difference between the maximum and minimum value may be 1.2 HRC. This corresponds to a standard deviation of ± 0.5 HRC. If the manufacturer is cautions, desiring to avoid disputes, he reduces the product tolerance of ± 2 HRC to ± 1.5 HRC so as to take into account the uncertainty of the tester. In this case (Fig. 29 b) already 14 % of heat treated pieces are to be rejected, in place of the 5 % mentioned earlier.

REFERENCES

(The numbering employed in the Bibliography "The metrology of hardness scales" - 1981 was retained in the present publication).

B-2 Barbato G., Desogus S. : Realizzazione di alcune scale di durezza presso l'Istituto di Metrologia "G. Colonnetti" di Torino. Ingegneria meccanica 1979, No. 4, 81-84.

B-5 Bochmann G., Hild K.: Der Einfluss von Abweichungen des Eindringkörpers von der Sollform auf die Härtewerte bei dem Messverfahren nach Rockwell C. Zeitschrift für Instrumentenkunde 68 (1960), 155-162.

B-10 Barbato G.: Investigations on the uncertainty of hardness measurement. Bulletin OIML No. 91 - Juin 1983, 3-12.

B-14 Barbato G., Desogus S., Li F., Weng Q.: Comparison of hardness measurements between China (NIM) and Italy (IMGC). Bulletin OIML No. 105, 1986, 18-23.

B-15 Barabanow S.P., Kotchin O.M.: Gosudarstvenniy spetsialniy etalon edinits tverdosti po shkalam Brinellya. Izmeritelnaya Tehnika, 1987, No. 5, 3-4.

B-16 Baker G.M., Yamamoto H., Yamamoto T.: Investigation of the indicated values of hardness blocks made in the USA and in Japan. Journal of material Testing Research Association of Japan, 33/1988/, No. 2, 111-116 (Available in English).

C-1 Čutka B.: Messeigenschaften der Rockwell-Härtenormalgeräte Typ Wazau (Halle). Feingerätetechnik 15 (1966), 569-575. (More in detail, in Czech language: Měrová Technika 7 (1968), 17-26).

C-3 Čutka B.: Die Genauigkeit der Härteprüfung in der tschechoslowakischen Industrie. Feingerätetechnik 19 (1970), 404-405.

C-6 Čutka B.: Metrology of Vickers hardness (in Czech language). Československá standardizace, 1984, 253-255.

C-7 Čutka B.: Metrology of Rockwell - A,B,C hardness (in Czech language). Československá standardizace, 1985, 159-164.

C-9 Čutka B.: The Czechoslovak State Standard of the Rockwell-N,T hardness scales (in Czech language). Československá standardizace, 1988, 187-191.

C-11 Čutka B.: Possibilities of the international unification of hardness scales (in Czech language). Accurate Measurement and Metrology of Hardness, ČSVTS, Praha, 1988, 96-116.

C-13 Čutka B: Hierarchy scheme of hardness measurement in Czechoslovakia (in Czech language). Měrová technika 1965, 96-103.

D-1 Dambacher H.: Fragen zur Normung der Härteprüfverfahren und Härteprüfgeräte nach Brinell, Rockwell und Vickers. VDI-Berichte Nr. 308, 1978, 77-90.

D-2 DeBellis A.: Maintaining Rockwell Hardness Standards. Materials Research and Standards, Vol. 12, No. 8, 1972, 28-30.

E-1 Eckardt H., Kersten W.: Stand der Makroeindruck-Härtemessung in der DDR-Industrie. Standardisierung und Qualität 20 (1974). 344-347.

E-4 Eckardt H., Kersten W.: Interaction between indenters and HRC standard machines. Personal communication on experiments performed in ASMW, Berlin in 1982.

F-1 Flurschütz F.: Einige Betrachtungen über Normal-Härteprüfgeräte und die Aufstellung einer Härtebezugsbasis. Feingerätetechnik 15 (1966), 564-569.

F-3 Frank K.: Die Bedeutung der Kontrollplättchen bei der Härteprüfung. Werkstattstechnik 31 (1937), 213.

G-1 Gidel A.Yu.: Method opredeleniya pogreshnostey formi nakonechnikov dlya tverdomerov. Ismeritelnaya technika 1967, No. 3, 33-38.

G-2 Gosset A.: Intercomparaison des échelles de dureté HRC et HV 30 entre la France et la Hongrie. Bulletin BNM, No. 42, octobre 1980, 38-43.

H-4 Hild K.: Der Einfluss der Toleranzen des Messverfahrens nach Rockwell C auf die Härtewerte. Zeitschrift für Instrumentenkunde 66 (1958), 202-207, 230-234.

H-6 Hild K.: Untersuchungen über den Einfluss der Gestalt von Eindringkörper und Probe in der Härtemessung nach Rockwell-C und Vickers. VDI-Berichte Nr. 41, 1961, 47-50.

H-7 Hormuth K.: Der Einfluss der Belastungszeit bei Rockwell-Härtemessungen hoher Genauigkeit. VDI-Berichte Nr. 41, 1961, 105-108.

H-8 Hormuth K.: Das Normalgerät des Deutschen Amtes für Messwesen zur Bestimmung der Vickershärte. Feingerätetechnik 11 (1962), 250-253.

H-9 Hormuth K.: Mesurages internationaux de comparaison de dureté et proposition pour la fixation d'une base de référence pour le procédé de mesurage de dureté Rockwell. Bulletin de l'OIML No. 14 (1963), 21-30.

H-10 Hormuth K.: Zum gegenwärtigen internationalen Stand auf den Gebieten der Kraft- und Härtemessung. Feingerätetechnik 15 (1966), 482-491.

H-20 Helmrath Ch., Gould D.: The development of a unified European HRC hardness scale (also in German). VDI-Bericht Nr. 804, 309-324.

I-1 Iizuka K., Imai H.: Form measurement of small spherical surfaces by means of the method of least squares. Bulletin of the Japan Soc. of Prec. Engg., Vol. 4, No. 1, 1-6.

I-2 Iizuka K., Yano H., Hida N., Goto M., Yamamoto K.: An international comparison of hardness conversion between Vickers and Rockwell C scales. Proceedings of the Round Table Discussion on Hardness Testing. 7th IMEKO, London, 1976, 89-105.

I-5 Iizuka K.: A comment on the uncertainty of hardness measurement. 9th IMEKO Congress, TC 5 Round Table, Berlin-West, 1982.

J-2 Jegaden G., Siener P., Voinchet J.: Die Prüfung der Härte von Blechen. VDI-Berichte Nr.100, 1967, 77-87.

K-14 Kotchin O.M., Koslow V.I.: O vvedenii etalonnoy shkali tverdosti C Rokvella. Ismeritelnaya tehnika, 1983, No. 3, 48-49.

K-18 Kotchin O.M.: Sovremennoye metrologitcheskoye obespetcheniye ismereniy tverdosti. In: Sovremenniye metrologitcheskiye problemi fisiko-technitcheskih ismereniy. Moskwa 1988.

K-21 Kovács G.: The effect of correction values at international hardness comparisons. Mérésügyi Közlemények XXVII (1986), 88-94.

M-1 Marriner R.S.: Performance tests on indentors for Rockwell hardness testing. Engineering, July 1, 1955, 17-19.

M-4 Marriner R.S., Wood J.G.: Developments in the accurate measurement of indentation hardness. Joint Convention on machines for materials and environmental testing, 6-10th Sept. 1965, paper 10, 79-84, Inst. mech. Engrs.

M-5 Marriner R.S.: Leistungsteste an Eindringkörpern, die der Rockwell-Härteprüfung dienen. Microtecnic XV (1961), 279-288.

M-9 Marriner R.S., Wood J.G.: Comparison of international Rockwell C and Vickers HV 30 hardness scales during 1970 and 1971. Bulletin of OIML, No 48, Sept. 1972, 19-31.

M-11 Meyer K.: Das Erreichen der Vergleichbarkeit der Härtemesswerte - ein problem der Werkstoff-Prüftechnik von internationaler Bedeutung. VDI-Berichte Bd.11, 1957, 103-122.

M-24 Meyer K.: Die Eingliederung des Härteprüfverfahrens nach Rockwell-C in das betriebliche Mess- und Prüfwesen. Werkstattstechnik, Nov. 1952.

M-26 Mikoszewski J.: Twardościomerze (Hardness testers. In Polish), Warszawa 1985.

O-3 OIML SP 19-Sr 4: Compte rendu de comparaison internationale des échelles de dureté. Prague 1984.

P-1 Petik F.: Diskussionsbeitrag. 1. Internationaler Erfahrungsaustausch Kraft- und Härtemessung, Berlin, 1966, 18-20.

P-2 Petik F.: Über die Genauigkeit des ungarischen Normal-Härteprüfgerätes nach Rockwell. VDI-Berichte Nr. 100, 1967, 157-159.

P-3 Petik F., Kovács G.: Experiments to ensure the reproducibility and conformity of national hardness scales. ACTA IMEKO 1967, Vol. I, 369-379.

P-4 Petik F., Čutka B., Hadjiski I., Kersten W., Mikoszewski J.: Vergleichsmessungen zwecks Vereinheitlichung der Härtebezugsskalen von fünf Ländern. Vorträge vom V. Kongress für Materialprüfung, Budapest 1970, Bd. I, 163-168.
(A more detailed version in Hungarian in: Finommechanika 11/1972/, 137-143).

P-5 Petik F,: Über einen Vorschlag für eine einheitliche Methode zur Festlegung der Grösse der Unsicherheit beim Übertragen der Härte-Besugsskalen mittels Härtevergleichsplatten. VDI-Berichte Nr. 160, 1972, 205-209.

P-6 Petik F., Marriner R.S.: Systematic differences of two standard scales of Rockwell C hardness as established by repeated comparisons. Bulletin of OIML No. 49, Dec. 1972, 21-25.

P-7 Petik F. Wechselwirkung zwischen Härtenormalgerät und Normaleindringkörpern bei Härtemessungen nach Rockwell C. Materialprüfung 15 (1973), 239-241.

P-9 Petik F.: On the choice of the confidence level for statistical tests of systematic differences. Proceedings of the Symposium on Measurement Theory, Enschede, Dec. 1975, 1-11.

P-10 Petik F.: Reliability of the determination of systematic differences between measurements by different Rockwell C indenters. Proceedings of the Round Table Discussion on Hardness Testing. 7th IMEKO, London 1976, 83-88.

P-11 Petik F.: Zur Metrologie der Härteskalen. Feingerätetechnik 26 (1977), 175-177.

P-12 Petik F.: Die staatliche Sicherung der Einheitlichkeit der Härtemessung. VDI-Berichte Nr. 308, 1978, 71-76.

P-13 Philipps C.E., Fenner A.J.: Analysis of the Rockwell hardness test, employing a new deadweight testing machine. Engineering, July 13, 1951, 57-60.

P-14 Piliptchuk B.I.: Pogreshnosti ismereniya tverdosti po Rockwellu. Trudi WNIIM 50 (110), Moskwa-Leningrad 1961, 5-21.

P-17 Petik F., Weiler W.: Hardness Standard Machines of National Institutes for Metrology. PTB-Me-32, Braunschweig, 1981 - Part B: 1983.

P-18 Petik F.: Problems of hardness measurement. Measurement 1 (1983), No. 1, 24-30.

P-20 Petik F.: Thesis presented at the Hungarian Academy of Science. Budapest, 1974.

P-21 Petik F.: Factors influencing hardness measurement. BIML, Paris 1983.

P-22 Petik F.: Hardness test blocks and indenters. BIML, Paris 1984.

P-23 Petik F.: Error propagation and traceability of hardness measurements. 9th IMEKO Congress, Round Table TC 5, Berlin-West, 1982.

P-24 Petik F.: Statement of uncertainties at hardness testing.
- Zur Angabe der Unsicherheit bei der Härteprüfung.
VDI-Berichte Nr. 583, 1986, 41-62.

P-25 Petik F.: Metrology of hardness. Past development and present state of the art. Measurement Vol. 8. No. 1, 1990, 42-44.

P-26 Petik F.: The effects of a change of loading cycle parameters in hardness testing. Bulletin OIML No. 113, Dec. 1988. 35-40.

P-27 Petik F.: Measured quantities uncertainties and interconnection between tolerances for hardness testers. VDI-Berichte Nr. 804, 325-346 (also in German).

P-28 Petik F.: Hardness standard equipment, BIML, Paris, 1989.

P-29 Petik F.: Hardness test blocks and their verification (in Hungarian), Mérés és Automatika II (1954), 120-122.

R-2 Ratiu M., Lambert V.: Beitrag zur Darstellung der Brinellhärteskale. Die Arbeiten der V-ten Konferenz für Schweisstechnik und Metallprüfung, Timisoara 1966, Ergänzungsband, 149-162, und Revue Roumaine des Sciences Techniques, Série de Métallurgie, Tome 10 (1965), 219-233.

R-12 Rossow E.: Über einige Vergleichsversuche mit der Härteprüfung nach Vickers. VDI-Berichte Nr. 41, 1961, 157-169.

R-13 Rossov E.: Präzision von Härtemessungen nach Rockwell-C mit Prüfgeräten handelsüblicher Bauart. VDI-Berichte Nr. 100, 1967, 29-34.

S-8 Smolitch S.A.: Voprosi edinstwa ismereniy tverdosti po shkalam Rockwella. Trudi WNIIM 91 (151), Moskwa-Leningrad 1967, 54-57.

S-11 Shin S., Hida N., Yano H.: Establishment of standard for Brinell hardness and evaluation of the error. Bulletin of NRLM 31 (1982), No. 4, 333-343.

S-18 Shinozuka K., Matsuura T.: The investigation of the accuracy of forty Rockwell hardness testing machines. (In Japanese). Report of the NRLM. Jan 1973, 51-60.

S-19 Shinozuka K.: The investigation of accuracy of hardness testers of 25 local industrial testing laboratories (In Japanese). Report of the NRLM. Oct 1975, 57-70.

T-1 Tolmon F.R., Wood J.G.: Precision measurement of Rockwell hardness diamond penetrators. Engineering, July 20, 1951, 89-90.

T-4 Tsai Cheng-ping, Li Fu-hsin and others: The design of the NIM Laser Type National Deadweight Standard Rockwell and Rockwell Superficial Hardness Machine. National Institute of Metrology, Pekin.

T-5 Tanaka K., Yano H.: Errors of visual judgement in precision measurements. Ergonomics, 1984, Vol. 27, 767-780.

W-2 Wartmann R.: Ermittlung der Genauigkeit und Empfindlichkeit verschiedener Härteprüfverfahren. Schweizer Archiv, Juli 1960, 269-278.

W-13 Weingraber H. Technische Härtemessung. München 1952.

W-23 Wood J.G.: Traceability to NPL for hardness testing. NPL News No. 355 (1982), 9-12.

Y-1 Yamamoto H.: Über die Entwicklung von Härtevergleichsplatten. VDI-Berichte Nr. 308, 1978, 179-183.

Y-3 Yamamoto K., Iwasaki M.: Measurements of apex profile of Rockwell diamond penetrators. Bulletin of the Central Inspection Institute of Weights and Measures, Tokyo, Series Number 3, Oct 1959, 20-27. The same in German; Messung des Spitzenprofils von Eindringkörpern nach Rockwell-C aus Diamant. VDI-Berichte, Nr.41, 1961, 123-127.

Y-4 Yamamoto K., Yano H., Yazima N.: Standardization of Rockwell hardness from a statistical viewpoint. Bulletin of the Central Inspection Institute of Weights and Measures, Tokyo, Series Number 3, Oct 1959, 9-19. The same in German: Über die Normung der Rockwell-Härte vom Standpunkt der Statistik aus betrachtet. VDI-Berichte, Nr. 41, 1961, 133-140.

Y-5 Yamamoto, Yano H.: Standardisierung und internationaler Vergleich der Rockwell-C-Härteskale, aufgestellt vom National Research Laboratory of Metrology, Tokyo, Feingerätetechnik 15 (1966), 507-513.

Y-6 Yamamoto K., Yano H.: Standardization of Rockwell C scale done by the National Research Laboratory of Metrology.
Bulletin of the NRLM, Series Number 13, Oct 1966, 1-9.

Y-8 Yamamoto D., Yano H.: Studies on determination of standard Rockwell diamond indenters. Bulletin of the NRLM, No. 14, 1967, 11-25, or Report for the meeting of WG Y-3, OIML, Oct. 1966, Wien.

Y-10 Yamashiro S., Uemura Y., Kawagoe H.: Auswirkungen von verschiedenen Fehlern auf den Härtemesswert nach Super-Rockwell. VDI-Berichte Nr. 160, 1972, 115-136.

Y-12 Yano H., Ishida H., Kamoshita T.: Characteristics of Standard Rockwell Diamond Indenters and methods of establishing Standard Indenters. Proceeding of the Round Table Discussion on Hardness Testing. 7th IMEKO, London, 1976, 16-50.

Y-13 Yoshizawa T.: Überblick über die Untersuchungen auf dem Gebiet der Härteprüfung in Japan. VDI-Berichte Nr. 41, 1961, 33-41.

Y-14 Yoshizawa T., Terasawa M.: Die Entwickung von Härtevergleichsblöcken in Japan. VDI-Berichte Nr. 100, 1967, 145-155.

Y-17 Yano H., Ishida H., Shin S.: Formalization and standard unification of Rockwell hardness test. 9th IMEKO Congress, TC 5 Round Table, Berlin-West, 1982.

Y-18 Yamamoto H., Yamamoto T.: The existing state of the accuracy of Rockwell hardness test. (In Japanese.) Journal of Material Testing Research Association of Japan, 29 (1984), 199-204.

Y-19 Yano H.: Personal differences in visual judgement in precision measurement. Proceedings ISMET 1978, Tokyo.

Y-20 Yamamoto K., Yano H.: International comparison of Rockwell C hardness done by the NRLM. Bulletin of the NRLM, Series Number 13, Oct. 1966, 10-31.

Y-26 Yano H., Kamoshita T., Yajima N., Yamamoto K.: The measured values of hardness by the local industrial testing laboratories and the makers of hardness tester. (In Japanese). Report of the NRLM, Nov 1972, 32-68.

STANDARD SPECIFICATIONS AND RECOMMENDATIONS

ISO 146-1984 Metallic materials - Hardness test - Verification of Vickers hardness testing machines HV 0.2 to HV 100.

ISO 156-1982 Metallic materials - Hardness test - Verification of Brinell hardness testing machines

ISO 716-1986 Metallic materials - Hardness test - Verification of Rockwell hardness testing machines (scales A-B-C-D-E-F-G-H-K)

ISO 1079-1989 Metallic materials - Hardness test - Verification of Rockwell superficial hardness testing mcahines (scales 15N, 30N, 45N, 15T, 30T and 45 T)

ISO 6508-1986 Metallic materials - Hardness test - Rockwell test (scales A-B-C-D-E-F-G-H-K)

VIM - International Vocabulary of Basic and General Terms in Metrology. BIPM - IEC - ISO - OIML, 1984

OIML International Document D 5-1982 Principles for establishement of hierarchy schemes for measuring instruments

TGL 31543/31 Staatliches Etalon der Einheit der Härte nach Brinell (Berlin)

GOST 8062-79 State standard and hierarchy scheme for hardness measurements on the Brinell scale (In Russian)

GOST 8063-79 State standard and hierarchy scheme for hardness measurements on the Vickers scale (In Russian)

GOST 8064-79 State standard and hierarchy scheme for hardness measurements on the Rockwell and Super-Rockwell scales (In Russian)

INDEX